

محركات البحث على الإنترنت

أ.م.د. زين الدين محمد عبد الهادي

أستاذ علم المعلومات المساعد

رئيس قسم المكتبات والمعلومات

كلية الآداب - جامعة حلوان

القاهرة

إيبيس. كوم للنشر والتوزيع

2007

محركات البحث على الإنترنت للمكتبات ومراكز المعلومات

د. زين عبد الهادي
رئيس قسم المكتبات والمعلومات
كلية الآداب - جامعة حلوان

القاهرة
إيبيس.كوم للنشر والتوزيع
2007

بسم الله الرحمن الرحيم

﴿ آمَنَ الرَّسُولُ بِمَا أُنزِلَ إِلَيْهِ مِن رَّبِّهِ وَالْمُؤْمِنُونَ كُلٌّ آمَنَ بِاللَّهِ وَمَلَائِكَتِهِ وَكُتُبِهِ وَرُسُلِهِ لَا نُفَرِّقُ بَيْنَ أَحَدٍ مِّن رُّسُلِهِ وَقَالُوا سَمِعْنَا وَأَطَعْنَا غُفْرَانَكَ رَبَّنَا وَإِلَيْكَ الْمَصِيرُ ﴾

صدق الله العظيم

المحتويات

الفصل الأول: محركات البحث: المفهوم والبنية والتاريخ:

- تعريف محركات البحث: -- 3
- بنية محرك البحث على الإنترنت: ------------------------------ 7
- أهمية محركات البحث: -- 12
- الأساليب المستخدمة للبحث: ----------------------------------- 13

الفصل الثاني: محرك بحث (الضوء الشمالي):

- خصائص المحرك: -- 18
- خصائص البحث في المحرك: ------------------------------------- 19
- السمات العامة للمحرك: --------------------------------------- 21
- نتائج البحث في المحرك: -------------------------------------- 22
- طرق البحث المستخدمة في نورثرن لايت: ------------------------ 23
- ميزات خاصة لمحرك نورثرن لايت: ------------------------------ 24

الفصل الثالث: محرك بحث (GOOGLE):

- خصائص المحرك: -- 28
- خصائص البحث في المحرك: ------------------------------------- 29

- السمات العامة للمحرك: -- 31

- نتائج البحث في المحرك: -- 34

- طرق البحث المستخدمة في جوجل:------------------------------ 35

الفصل الرابع: محرك بحث (ALLTHEWEB):

- خصائص المحرك: -- 39

- خصائص البحث في المحرك: ------------------------------------ 41

- السمات العامة للمحرك:-- 43

- نتائج البحث في المحرك: -- 44

- طرق البحث المستخدمة في فاست: -------------------------------- 45

الفصل الخامس: محرك بحث (LYCOS):

- خصائص المحرك: -- 49

- خصائص البحث في المحرك:------------------------------------ 49

- السمات العامة للمحرك:-- 50

- نتائج البحث في المحرك: -- 51

الفصل السادس: محرك بحث (EXCITE):

- خصائص المحرك: -- 55

- خصائص البحث في المحرك:------------------------------------ 56

- السمات العامة للمحرك: -- 58

- نتائج البحث في المحرك: -- 58

الفصل السابع: محرك بحث (ALTAVISTA):

- خصائص البحث: -- 61

- خصائص البحث في المحرك: -- 62

- السمات العامة للمحرك: -- 62

- نتائج البحث في المحرك: -- 63

- قياس أداء المحركات المختارة: -- 64

- نتائج البحث المقارن بين المحركات الست: -- 72

- نتائج دراسة محركات البحث السالفة: -- 84

- النتائج الخاصة بكل محرك: -- 90

- النتائج المقارنة بين المحركات: -- 99

مــقــدمــة

لعله من نافلة القول أن الحياة بلا هواء لا يمكن أن تكون حياة، ومحركات البحث بالنسبة لشبكة الإنترنت هى الهواء والماء، ولا يمكن استخدام شبكة الإنترنت بدون محركات البحث، إذ يعد الأمر في تلك الحالة كالبحث عن إبرة في كومة قش.

ثمــة تغيرات هائلة حدثت في أدوات البحث على الإنترنت خلال السنوات الأخــيرة، هــذه الأدوات تنقســم إلــى أقســام أربــع هــى أدلة البحث Search Directories ومحــركات البحث Search Engines وما وراء محركات البحث Meta Search Engines وأدوات البحث الذكية Intelligent Search Agents وتــتفرع عــن هــذه الأقســام فروع أخرى تتوجه بالبحث إلى فئات بحث أخرى كمحــركات البحث المتخصصة في قضايا معينة، أو محركات بحث خاصة بفئات عمــرية أو نوعــية محددة، وتستمر هذه الانقسامات طالما وجدت شبكة الإنترنت وطالما تتطور يوماً بعد آخر.

لقــد أخــذت علــى عاتقي في هذا العمل أن أتناول قسماً واحداً من أدوات البحث على شبكة الإنترنت هى محركات البحث من حيث مفاهيمها وأنواعها وأهم

المحـــركات العاملـــة وخصائصـــها وأهم الفروقات بينها، وقد أفردت جانباً لبعض القياســـات الإحصـــائية لقدرات هذه المحركات، ولا شك أن التطورات التي تحدث يومياً في شبكة الإنترنت وفي محركات البحث على وجه التحديد سوف تترك آثاراً بعيدة المدى على الشركات التي تقف خلف هذه المحركات وعلى الأداء الاقتصادي لهـــا بشـــكل عـــام، إضافة إلى أن هذه التطورات لا تكاد تتوقف مما يعني أهمية ملاحقتها ورصدها وتسجيلها.

هـــذا العمـــل علـــى صغر حجمه إلاَّ أنه موجه لكل شخص يستخدم شبكة الإنترنـــت وإلى الطلبة دارسي علوم المعلومات وشبكة الإنترنت، وإلى المسئولين عن إدارات وأقسام المعلومات، وإلى أخصائي المكتبات والمعلومات وإلى كل مهتم بقضايا شبكة الإنترنت.

والله الموفق والمستعان،،

د. زيـــن عبد الهادي

الفصـــل الأول

محركات البحث
المفهوم والبنية والتاريخ

الفصل الأول

محركات البحث: المفهوم والبنية والتاريخ

مقدمة:

لماذا محركات البحث على الإنترنت؟

لكــل واحد منا رغبته التى تقوده عند البحث على الإنترنت، فقد نبحث عن مواصفات سيارة، أو عن حياة إنسان ما مثل نيل أرمسترونج الذي سار فوق سطح القمــر، أو عن شراء تذاكر للسينما، أو لحجز موعد عند طبيب أو للحصول على شــهادة أكاديمية، والبعض منا يجرى بحوثه الأكاديمية باستخدام هذه الأدوات عبر الإنترنت، أو يقوم بإجراء بحث للتعرف على اتجاهات السوق المتعلقة بسلعة معينة يريد ترويجها.

لكن هل يعرف كل منا كيف يستخدم هذه الأدوات بشكل دقيق وموضوعي، وبحيث نحصل على ما نريد بأقل جهد وفى أسرع وقت؟ هل هناك محركات بحث متخصصــة فى موضوعات وقضايا محددة؟ وما الفرق بين أدوات البحث المختلفة على الإنترنت.

هــذا العمــل هدفــه الــتعريف ببعض أدوات البحث على الإنترنت وهى محــركات البحث عبر فهم بنيتها وتركيبها وفوائدها ومن ثم التعرض لبعض منها والتعرف على خصائصه والتي يمكن أن تساعدنا كآباء وأمهات ومتخصصين وغير متخصصين على التجول على الإنترنت بحرية أكبر، والحصول على ما نريد.

وإذا كان مشوار الألف ميل يبدأ بخطوة، فإن خطوتنا الأولى فى هذا الفصل هو التعرف على البنية التركيبية لمحركات البحث ومفهومها واستخداماتها وأساليب البحث المستخدمة.

تعريف محركات البحث:

محــركات البحــث هى أدوات بحث تعمل من خلال استراتيجيات محددة (البحث البوليني) أو استراتيجيات مفتوحة (البحث باللغة الطبيعية) وذلك للبحث في حقــول Fields أو وثائق نصية Text Documents، والأكثر من ذلك أنها يمكن أن تبحث عن أشياء Objects (صور ورسوم وخرائط وأصوات)، في بيئة محددة هــى بيــئة شــبكة الإنترنت، وذلك يعني أنها تبحث في ملايين المواقع ومليارات الكلمــات فــي وقت محدد وتتميز بسرعة الاستجابة، وعادة ما تكون إجاباتها إما

مواقـــع على الإنترنت تتوافر فيها كل المصطلحات التي تم البحث عنها أو بعضها، أو مواقـــع محـــددة ســـلفاً مـــن خـــلال ما يعرف بقائمة أو دليل البحث Search Directory.

كذلـــك فـــإن محـــركات البحـــث تعتبر النتاج الطبيعي لتطور المتصفحات Browsers على شبكة الإنترنت، فقد كان لابد من توافر وسائل وأدوات تمكن من البحـــث على مواقع الإنترنت بسرعة وسهولة، وفي ذات الوقت بشكل غير معقد، وأن يتآلف معها عدد كبير من المتعاملين على الشبكة.

وقـــد تطورت هذه المحركات بحيث أصبحت تقام عليها مؤسسات بأكملها تـــبلغ استثماراتها ملايين الدولارات وتشهد هذه المؤسسات العديد من الظواهر التي ترتـــبط باقتصـــاديات الســـوق كالشراكة Partnership والتحالفات الاستراتيجية والإندماجات وغيرها من ظواهر الصناعة الحديثة المرتبطة بعالم المعلومات.

وقد ارتبط ظهور محركات البحث بظهور متصفحات شبكة الإنترنت، كما ارتـــبط بمعضـــلة كبيرة كانت تواجه الباحثين من قبل، هذه المعضلة تشبه إلى حد كبير معضلة البحث في فهارس مكتبة عملاقة، فكيف يكون الحال إذا كنا نبحث عن عـــنوان كتاب معين في ذات الوقت الذي لا يتوافر فيه فهرس عنوان بالمكتبة، أو

كيف يكون الموقف عند الرغبة في البحث عن الأوعية والكتب التي تناولت قضية معينة أو موضوعا معينا حيث لا يتوافر فهرس للموضوعات في المكتبة، وبالطبع فالأمـــر أكثر خطورة على شبكة الإنترنت فأنت تبحث في ملايين المواقع عن شئ محدد، أو كما يقول المثل العامي "البحث عن إبرة في كوم من القش".

لقـــد أصبح الموقف أعقد من ذلك، فأحياناً نضطر للبحث عن شخص معين على الإنترنت، عن رقم هاتف معين، عن بلدة محددة، إنها أسئلة واستفسارات من النوع البسيط، لكنها في ظل عدم وجود أدوات تنظم هذا الكم الرهيب من الصفحات علـــى الإنترنت، فإن البحث في هذه الحالة يصبح نوعاً من العبث الذي لاطائل من وراءه.

لقـــد استغرق بناء هذه المحركات العديد من السنوات على شبكة الإنترنت، ومـــن الناحية الفنية فهي برامج آلية صممت للعمل في بيئة الإنترنت، سميت باسم محـــركات البحث أو أدوات البحث وهي في الحقيقة "نظم لإدارة قواعد البيانات تم تطويـــرها لإدارة البيانات وقد طورت بصفة خاصة للعمل على الإنترنت [1] وهى وســـيلة البحث عن المعلومات على الإنترنت، وعلى ذلك فإن مصممي متصفحات

[1] Platt, Nina. Website search engines. November 1996. the piper letter: Databases, Features and special reports. www.piperinfo.com/p103/search.html
Last updated: Thu. Feb. 20,1997 retrieved and printed: 22-12-1998

الـــوب على الإنترنت قد طوروا مواجهات للوب تمكن مستخدمي المتصفحات من الاتصال بمحركات البحث، أي أنه من الناحية الفنية لكي يتم البحث على الإنترنت فـــإن هناك ثلاثة أنواع من الاتصالات تتم كي يمكن استخدام الأدوات المتاحة في محـــرك البـــحث حـــتى نتمكن من البحث على الإنترنت، هـــذه الاتصالات تسير كالتالي :

شكل (1) طرق الاتصالات على شبكة الإنترنت للوصول إلى محرك البحث

الحاسب	← مورد خدمة الإنترنت	← متصفح الإنترنت	← محرك البحث
	اتصال هاتفي (قد لايتم في حالة وجود خط مؤجر Leased Line مفتوح بصفة دائمة على الإنترنت)	برنامج محمل على الحاسب ينطلق حالما تم الاتصال	قد تكون صفحة أولى Home Page أو تستخدم كأداة من صفحة أولى لموقع على الإنترنت مثل مايكروسوفت أو قناة CNN للأخبار

2- بنية محرك البحث على الإنترنت Structure of Search Engine :

لقد سبقت الإشارة إلى أن محرك البحث عبارة عن برنامج Program أو نظام لإدارة قواعد البيانات، هذا النظام مكون من ثلاثة أجزاء تعمل في تناسق بهدف التمكين من البحث على الإنترنت، هذه الأجزاء هى :

1- الزاحف أو العنكبوت Crawler or Spider

وهو جزء من البرنامج ينطلق لزيارة صفحات الشبكة العنكبوتية على فترات منتظمة متكررة قد تكون نصف شهرية أو شهرية، وهو يتتبع الصفحات ذات الصلة (المربوطة) Linked بالموقع الأم ويتتبع كذلك أي تغيرات يمكن أن تكون قد تمت، وبالتالي فهو يمكننا من ربط كل الصفحات الفرعية بالصفحة الأم للموقع، ومن ثم ننتقل إلى الجزء الثاني من برنامج محرك البحث وهو الكشاف أو الفهرس.

2- الفهرس أو الكشاف Catalog or index

كل مرة يتم فيها زيارة أي موقع على الإنترنت من قبل الزاحف أو العنكبوت يقوم فيها هذا الأخير ببناء قائمة شبه آلية .. يجمع المصطلحات التي

وردت فـــي الصفحات التي تمت زيارتها، وعلى ذلك فهذه القائمة هى أشبه بمكنز آلـــي (قـــاموس ذو بـــناء معين) يتم تحديثه في كل مرة يقوم فيها الزاحف بزيارة المواقع على الإنترنت، وإذا حدث أي تغيير بين الزيارة الأولى والثانية للزاحف أو أنشـــأت صـــفحات جديـــدة فإن هذه التغييرات لن تكون محسوسة (لن يتم تكشيفها وفهرســـتها) إلا مـــع انـــتهاء الزيارة الثانية للموقع، حيث يتم إضافة المصطلحات الـــواردة بهـــذه الصـــفحات الجديـــدة أو التـــي تم تغييرها إلى هذا الكشاف (قائمة بالمصطلحات المستخرجة من المواقع مرتبة هجائيا) أو هذا الفهرس. ومن ثم ننتقل إلى الجزء الثالث من محرك البحث.

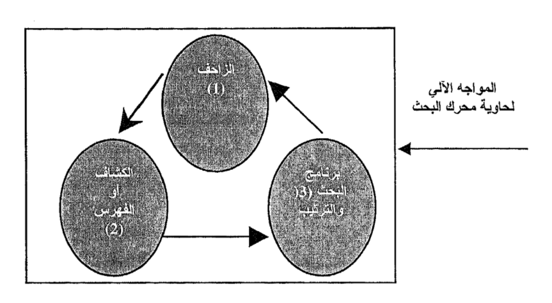
3- برنامج محرك البحث Search Engine Software

وهـــو الـــذي يقـــوم بالبحث في الكشاف الذي تم إنشاؤه للبحث بين ملايين الصـــفحات المســـجلة في الكشاف لكي يعثر على المواقع المتطابقة مع ما تم طلبه وأيضاً يقوم بوضع هذه الصفحات في شكل طبقي Ranking .[1]

---

[1] how search engines work. In web site titled: search engine watch on: www.searchenginewatch.com /webmasters/work.html. Retrieved & printed: 18/10/1998.

ويمكن من خلال الشكل التالي التعرف على هذه الأجزاء الثلاثة :



(شكل 2) بنية محرك البحث على شبكة الإنترنت

أسباب اختلاف النتائج من محرك بحث إلى آخر:

لكن يبقى السؤال، إذا كانت بنية محركات البحث المتوافرة على الإنترنت واحدة، فما هو السبب وراء اختلاف النتائج من محرك إلى آخر !!

وللإجابة على هذا السؤال، يمكن القول بأن السبب الرئيسي في ذلك يعود إلى مجموعة من العناصر يمكن الإشارة إليها بشكل مبدئي فى هذه المقدمة فيما يلي:

لاحظ: إن الصفحة الأم Homepage هو مصطلح مخادع تماماً، فالصفحة التي يفتح عليها متصفح الإنترنت عند تشغيله هي صفحة أم، كما أن قيامك ببناء موقع لنفسك على الإنترنت يمكن أن يطلق عليه أيضاً صفحة أم، وعلى ذلك فهي في نفس الوقت تعبر عن مصطلح موقع عنكبوتي website، وهي ذات الوقت فهي بالنسبة للآخرين موقع وليس صفحة أم.

1- الموقع Location كمكان مخصص للموقع على الإنترنت.

2- مدى تردد Frequency الزاحف على زيارة الموقع.

3- قد تكون هناك سياسة لعرض المواقع وذلك من قبل الشركة المالكة للمحرك (هذه السياسة غير معلنة فى أغلب الأحايين لعامة مستخدمي المحرك).

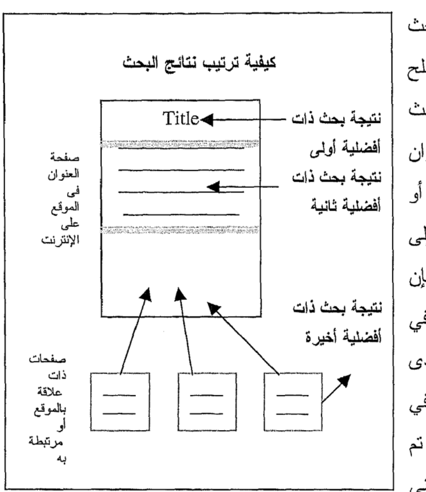
كما يمكن الإشارة إلى أن العنصر الأول هو الأكثر أهمية وتأثيراً في قضية عرض الإشارات للمواقع كنتيجة للبحث في موضوع معين أو في مواقع معينة في أحد محركات البحث المتناثرة على الإنترنت.

والـــذي يحـــدث هنا أننا حين نبحث – فى المكتبات على سبيل المثال–عن كلمة مفتاحية ما، فإننا نبحث أولاً في العنوان الخاص بالكتب أو المقالات، وهذا هو مـــا يفعله محرك البحث بالضبط إذ يعتمد مدى ترتيب الموقع بين أقرانه على مدى ظهور الكلمة المبحوث عنها في عنوان الموقـــع.(1)

---

(1) How search engines rank web pages. Ibid. p.1.

على ذلك فإن الكلمة التي ظهرت في عنوان موقع ما تضع هذا الموقع في مقدمة المواقع التي تم البحث فيها والعثور على المصطلح في العنوان، يلي ذلك البحث في أقرب جزء من عنوان الصفحة، كأقرب فقرة أو أقرب مجموعة فقرات إلى العنوان، وعلى ذلك فإن ترتيب نتائج البحث في المواقع يعتمد على مدى قرب أو بعد المصطلح في الصفحة أو الموقع الذي تم البحث فيه والعنصر الثاني



المؤثر في ترتيب الموقع في نتائج البحث هو مدى تردد وظهور المصطلح الذي تم البحث عنه بالنسبة لترددات المصطلحات الأخرى في صفحات الموقع، ويمكن أن نطلق على هذين العنصرين مجتمعين منهج التردد على الموقع Location /Frequency method.[1]

[1] Ibid

وبصـــفة عامة فإن متوسط عدد الصفحات التي يقوم أي محرك بحث على الإنترنـــت بفحصها دورياً يتراوح ما بين 3-15 مليون صفحة (ولا نقول موقع)، وتـــتراوح فترات تحديث البيانات التي يقوم المحرك بجمعها عن كل موقع ما بين دقـــائق إلـــى شهرين وربما كان محرك البحث إنفوسيك Infoseek (الباحث عن المعلومات) وقد كان هو المحرك الوحيد صاحب أقصر مدة تحديث للبيانات.

أهمية محركات البحث:

ربمـــا تعـــود محـــركات البحث في أهميتها إلى إنها تكاد تكون أهم وسيلة للبحـــث على شبكة الإنترنت، ولكن يمكن القول بشكل عام بأن عملية البحث على الإنترنـــت أقـــرب في مضمونها إلى البحث في المكتبات عن معلومات معينة، فإذا ســـلمنا بذلـــك فإن استخدام استراتيجيات بحث معينة وتوسيعها وتضييقها، يضعنا جميعا فى خندق واحد يتطلب منا الدقة فى اختيار المحرك الذي سنبدأ البحث منه.

يتطلـــب هـــذا الوضـــع تأهيل هؤلاء المتعاملين مع المعلومات للتعامل مع محـــركات البحـــث، ومـــن هـــنا تأتـــي أهمية محركات البحث بالنسبة لمستخدمي الإنترنت.

وعلى ذلك يمكن القول بأن أهمية محركات البحث تكمن في:

1- توسيع دائرة المؤسسات التي يمكن أن تجذب خريجي الجامعات للعمل بها خارج إطار المؤسسات التقليدية.

2- تلبية احتياجات المستفيد من المعلومات في المجتمع الافتراضي بالتعرف على سمات وخصائص هذه المحركات.

3- التعرف على الخصائص البحثية لكل محرك وبالتالي معرفة متى يمكن استخدام محرك معين في موقف معين.

4- تطوير المقررات الدراسية بالجامعات بحيث تحتوي على الموضوعات ذات العلاقة بالإنترنت لتلبي الموقف العالمي الجديد.

5- الحصول على إجابات سريعة وواضحة على كل الأسئلة التى يمكن أن يتوجه بها أي شخص إلى الإنترنت وبالتحديد إلى محرك بحث بعينه.

الأساليب المستخدمة للبحث:

تعتمد محركات البحث على مجموعة محددة من أساليب البحث فيها يمكن إجمالها فيما يلي:

1- استخدام اللغة الطبيعية Natural Query language

2- استخدام المعاملات البوليانية Boolean Operators

3- استخدام معاملات التقريب Proximity operators

4- البحث باستخدام الجمل	Phrase Searching
5- البحث باستخدام مكانز	Thesaurus
6- البحث باستخدام المفهوم	Concept searching
7- البحث باستخدام البتر	Truncation
8- البحث بأسلوب المطابقة الكاملة	Exact match
9- البحث بأسلوب المطابقة الاحتمالية	Fuzzy match
10- استخدام المعاملات الرقمية	Numeric operators
11- استخدام معامل المدى	Range operators
12- البحث في حقول محددة	Fielded searches
13- البحث باستخدام أمثلة	Query by examples
14- استخدام الإرشادات التي تظهر مع البحث في إعادة البحث	Advisors

إضافة إلى ذلك تتمتع بعض المحركات بخصائص بحث محددة، مما يعني أن لكل محرك بحث على الإنترنت شخصيته المتفردة التي يجب الإلمام بها، حيث يصعب أن يقدم محرك بحث على الإنترنت نفس النتيجة وهناك العديد من الدراسات التي يمكن الرجوع إليها في هذا الشأن.[1]

[1] Nueva Library help. Choose the best engine for your purpose.

الفصـــل الثانـــي

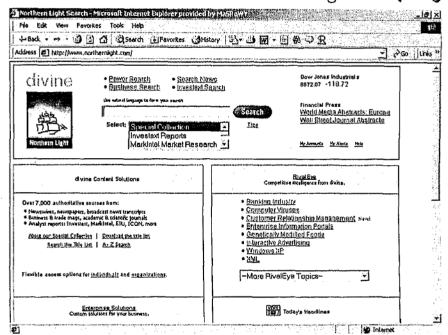
محرك بحث

"الضوء الشمالي"

الفصل الثاني

محرك بحث "الضوء الشمالي"

- محرك البحث: *Northernlight*



الصفحة الأم فى محرك نورثرن لايت

http://www.northernlight.com/

تقــول الشــركة المالكة لهذا المحرك على موقعها على شبكة الإنترنت أن هدفها تكشيف وتصنيف كل المعرفة البشرية بمعيار منطقي موحد. وقد بدأت العمل في سبتمبر 1995 في كامبريدج بولاية ماساشوستس، وتم إطلاق هذا المحرك في أغسطس 1997، من خلال صفحة واحدة وصلت الآن إلى 21 مليون صفحة على الــويب، إضافة إلى 7100 عنوان لمصدر معلومات موجودة فيما يعرف بمجموعة

الشـــركة الخاصة. وتحتوى هذه المجموعة من المصادر على مجموعة أخرى من الوثائق تصل إلى 25 مليون وثيقة نصية كاملة، تتراوح بين مقالات الصحف ذات النص الكامل والكتب والدوريات، وشبكات الأخبار، والمصادر المرجعية.

تقـــوم الشركة بترتيب المواقع الناتجة عن البحث وفقا لعنصرين هما عدد مرات ظهور الكلمة فى الصفحة أو الموقع، وأيضا ظهور الكلمة فى عنوان الموقع[1]. كما أنه لا يستخدم تركيبة *MetaTag* فى الملخص الذي يعرضه عن كل موقع كنتـــيجة. وتركيـــبة *Meta Tag* هى تركيبة الهدف منها وضع الكلمات المفتاحية الموضـــوعية والواصفات والأسماء ونوعية المحتوى داخل الصفحة بهدف أن يقوم محـــرك البحـــث بتكشيفها، وغالبا ما تعبر عن المحتوى الموضوعي للموقع ويقوم بإعدادها مصمم الصفحة على الوب*.

---

(1) Search Engine Tutorial for Web Designers: Northern Light.
url http://northernwebs.com/set/northern_lights.html> retrieved and printed: 2/9/2001

*تأخذ تركيبة Meta Tag هذا الشكل غالبا:

```
<Head>
<Ti
<Meta http-equiv=content-                     -
              مكتبة، مكتبات، تطوير المجموعات، المكتبات الرقمية
                   تطوير المكتبات فى العالم العربى
              محمد على
</Head>
```

إضافة إلى ذلك يمكن القول بأن هذا المحرك يعتبر واحدا من أكبر محركات البحث على شبكة الإنترنت حيث يقوم بتكشيف وفهرسة حوالي 140 مليون موقع على شبكة الإنترنت[1]، بمعنى أنه يقوم بتجميع الكلمات ذات الدلالة الموضوعية من كل موقع على الإنترنت قام بالعثور عليه، ويقوم بترتيبها ومن ثم البحث فيها بعد ذلك بناء على رغبة المستفيد.

خصائص المحرك:

يقوم هذا المحرك بالبحث فى مجموعة من المواقع التى تم ترتيبها وفق مجموعة من الفئات كالتالي:

1- المجموعة الخاصة Special collection
2- تقارير الاستثمار Investext Reports
3- تقارير وأبحاث MarkIntel
4- أرشيف الأخبار Archived News
5- البنوك الأمريكية American Banker
6- مجلة أسبوع الأعمال Business Week

(1) Ibid.

7- الاقتصادي Economist

8- مستخلصات جريدة وول ستريت جورنال

9- مستخلصات وورلد ميديا: أوروبا

والمشـــكلة هـــنا أيها تختار من هذه القطاعات لتقوم بالبحث فيه، حيث أن البحث يتم بشكل مبدئي على المجموعة الأولى منها أي المجموعة الخاصة، كما أن مشـــكلة هـــذا المحرك الأخرى أنه لا يبحث فى مواقع تم فهرستها وتصنيفها وفقا لـــتدخل الحاسب وإنما هذه المواقع تم تصنيفها بتدخل بشرى، إضافة إلى أن جميع النـــتائج الـــتى يوردها تتم بمقابل مادي يجب دفعه للحصول على أي من مصادر المعلومات.

خصائص البحث فى المحرك:

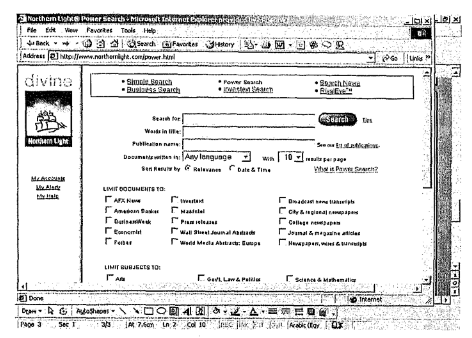
يمكن البحث فى المحرك باستخدام أربعة أساليب هى:

1- البحث البسيط

2- البحث المتقدم

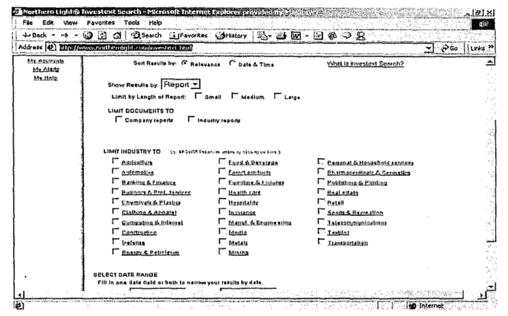
3- البحث عن ما يتعلق بإدارة الأعمال

4- البحث فى مجال الاستثمار (عن الشركات وعلاماتها التجارية واستثماراتها وتقاريرها الاقتصادية)



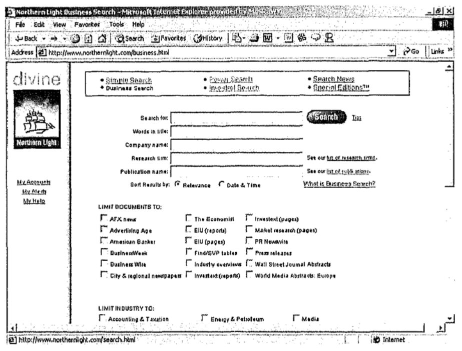
شكل يعرض للبحث المتقدم فى محرك (نورثرن لايت)

http://www.northernlight.com/power.html



شكل يعرض للبحث الاستثماري فى محرك (نورثرن لايت)

http://www.northernlight.com/investext.html



شكل يعرض للبحث فيما يتعلق بإدارة الأعمال فى محرك (نورثرن لايت)
http://www.northernlight.com/business.html

السمات العامة للمحرك:

يتميز هذا المحرك بقدراته الهائلة فى الأبحاث المتعلقة بالاقتصاد وإدارة الأعمال على شبكة الإنترنت، وعلى الرغم من أنه كان محركا عاما فى بداياته، إلا أنه مع الوقت أصبح متخصصا فى هذه القضايا وما شابهها.

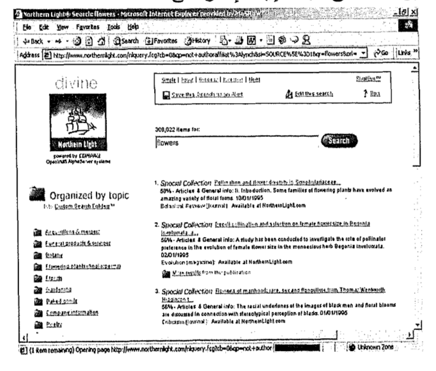
كما يتميز بقدراته البحثية الفائقة، حيث اعتمد على عمليات التكشيف والفهرسة اليدوية البعيدة عن الآلة أو ما يعرف فى علم المعلومات (بالتكشيف الآلي) Automatic Indexing أي أن الروبوت المتوافر فى محرك البحث هو الذي يقوم بتحديد المصطلحات التى يمكن البحث بها داخل المحرك عبر اقتناصه للكلمات ذات

الدلالة الموضوعية التي قام مسئول صفحة الوب ببثها داخل تركيبة الميتاتاج، أو أن يقوم المحرك بتكشيف كل أو بعض صفحات الموقع.

نتائج البحث في المحرك:

تختلف نتائج البحث من فئة لأخرى في هذا المحرك وفقا لنوع الفئة التي سنقوم بالبحث فيها، وعلى سبيل المثال لو قمنا بالبحث عن كلمة زهور flowers في هذا المحرك في المجموعة الخاصة special collection لكانت النتيجة كالتالي:

وجود **308,022** مادة في موضوع الزهور في قاعدة بيانات المحرك يقوم بترتيبها وفق نظام معين على بناء سبقت إليه الإشارة في المقدمة.



طرق البحث المستخدمة فى نورثرن لايت:

يمكـــن البحـــث فـــى نورثرن لايت بمجموعة مختلفة من الأساليب، يمكن الإشارة لها كالتالي:

1- استخدام اللغة الطبيعية فى البحث، مثال ماهى درجة غليان الماء؟

2- اســـتخدام المفـــردات البسيطة فى البحث مثل القاهرة أو مسجد عمرو بن العاص.

3- استخدام ما يعرف بالمعاملات البوليانية أو الجبرية Boolean Operators وهى المعروفة فى علم المعلومات بالمعاملات المنطقية (و – أو – ليس) أو (And-Or-Not)

4- البحث بجمل محددة وهنا تقوم بوضع الجملة بين علامتي تنصيص كالتالي "أهرام الجيزة" أو "أطلال أم كلثوم"

5- اســـتخدام أســـلوب البتر عند البحث، فعلى سبيل المثال إذا كنت تبحث عن كلمـــة معلومـــات وجميع اشتقاقاتها يمكنك كتابة "معلوم*" ويقوم المحرك باستحضار كل المواقع التى تحتوى كلمات مثل "معلومات ومعلومة ومعلوم ومعلومية معلوميات" وغيرها

6- البحث فى حقول محددة، وحيث أن المحرك يقوم بالبحث فى مجموعة من قواعد البيانات وفقا للفئات التى ذكرناها فإنه قد يبحث فى بعض حقول هذه القواعد مثل، عنوان الوثيقة أو اسم الشركة، أو النص داخل الوثيقة.

هـذه هـى الوسائل الست التى يدعمها هذا المحرك من واقع الوسائل التى ذكرناها فى الحلقة الماضية.

ميزات خاصة لمحرك نورثرن لايت:

أيضـا مـن المهـم الإشارة إلى أن هذا المحرك يوفر خدمة تسمى خدمة "الإحاطـة الجاريـة" Current Awareness بمعـنى أنه يمكنك تسجيل بياناتك والموضوعات التى تقع فى دائرة اهتمامك الشخصي ومن ثم يقوم المحرك بإمدادك كـل فـترة بالمواقع أو الوثائق التى يمكن أن تقع فى دائرة هذا الاهتمام، وهى من الخدمات التى يمتاز بها هذا المحرك دون غيره من المحركات.

يتـبقى أن أشـير إلى أن العرض الذي سوف نقدمه للمحركات الأخرى ثم المقارنـات التى سنجريها بين القدرات البحثية لهذه المحركات هى التى سوف تبين مـدى مـا يتمتع به كل محرك من قوة، وبالتالي نتعرض لمواقف القوة والضعف وحالات استخدام كل محرك من هذه المحركات.

الفصـــل الثالـــث

محرك بحث

" Google "

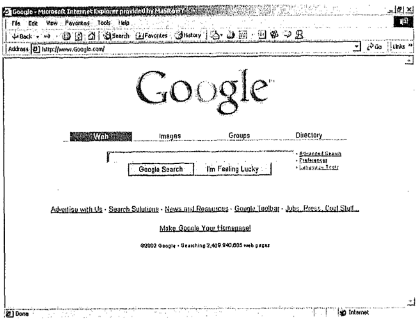
الفصل الثالث

محرك بحث " *Google* "

- محرك البحث: *Google*

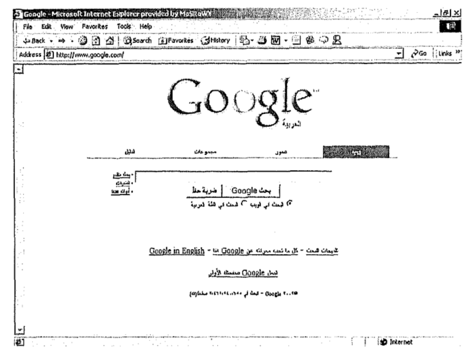
تقــول الشـــركة المالكـــة لهذا المحرك أن هدفها هو تنظيم معلومات العالم والوصـــول إلـــيها من أي مكان وبالتالي توسيع قاعدة الإفادة منها. لقد تم تأسيس الشركة المالكة لهذا المحرك عام 1998 بواسطة لاري بايج وسيرغي برين، وهما اثنان من طلاب الدكتوراه في جامعة ستانفورد. وقد أعلنت هذه الشركة الخاصة في حزيـــران/يونـــيو 1999 أنها حصلت على تمويل بقيمة 25 مليون دولار أميركي. ويشـــمل شركاء الشركة في التمويل Kleiner Perkins Caufield & Byers و Sequoia Capital.، كما تشير بيانات الشركة أيضاً إلى أنه باستخدام هذا المحرك يمكن البحث في 1.346.966.000 مليون صفحة على الإنترنت.

وربما ينبغي الإشارة هنا إلى أن معنى كلمة جوجل Google فهي تلاعب علـــى كلمـــة googol، التـــي اخترعها ملتون سيروتا، ابن أخت عالم الرياضيات الأميركـــي إدوارد كاسنر، للدلالة على رقم 1 تتبعه 100 صفر. ويعكس استخدام Googleلهـــذه الكلمـــة إصـــرار الشـــركة المالكة على تنظيم الكمية الهائلة من المعلومات المتوفرة على الشبكة وفي العالم.



شكل صفحة جوجل الإنجليزية على شبكة الإنترنت

www.google.com



شكل صفحة جوجل العربية على شبكة الإنترنت

http://www.google.com/language_tools?hl=ar

خصائص المحرك:

يقوم هذا المحرك بالبحث فى أربع فئات متوافرة على شبكة الإنترنت كالتالي:

1- الشبكة العنكبوتية Web

2- الصور Images

3- مجموعــات Groups (يغطــى هذا المحرك نحو 700 مليون رسالة تم تبادلها من خلال مجموعة يوزنت Usenet خلال العشرين سنة الماضية) .

4- دليل جوجل Directory (دليل بستة عشر موضوعا عريضا يمكن البحث فيها وفقا لما ترغب به من هذه الموضوعات)

ويتم البحث هنا مباشرة على المجموعة الأولى من الخيارات الأربعة وهى الشبكة العنكبوتية، إلا إذا أردت البحث فى مجموعة أخرى من المجموعات الثلاث الباقية.

وربما يجب الإشارة هنا إلى مفارقة هامة يجب أن يحتاط لها كل من يقوم بالبحث على شبكة الإنترنت، وهو أن **هناك ما يعرف بدليل أو شجرة البحث، وذلك بجانــب محرك البحث، بمعنى أن المجموعة الرابعة من جوجل تمثل وسيلة فعالة**

للمبتدئين بالبحث على شبكة الإنترنت، وتكاد هذه المجموعة تمثل قاسما مشتركا أعظــم بين كل أدوات البحث على شبكة الإنترنت، حيث يمكن للراغب استخدامها كوسيلة أولى للبحث على الإنترنت مع ما فى ذلك من مخاطر، فالمواقع المذكورة فـــى أدلـــة البحث عادة تخضع لاختيار الشركة صاحبة الموقع أو الدليل وبالتالي فمعايـــير اختـــيارها لا يمكـــن الحكم عليها بسهولة فى هذا المقام، وهو ما ينبغي الإشارة إليه هنا.

خصائص البحث فى المحرك:

ربمـــا ينبغـــي الإشارة أيضا إلى أن **جوجل كمحرك بحث يعتبر محركا ذو صيغة عالمية حيث يوفر واجهة استخدام بعدد 79 لغة منها العربية،** وهو ما قد لا يتوافر فى محركات بحث أخرى على شبكة الإنترنت.

يدعم جوجل البحث البسيط والبحث المركب، بمعنى أنه يمكنك أن تكتب ما تريد فى مربع البحث Search Frame وسوف يقوم بتنقية البحث الخاص بك، إن ماكتبـــته هـــو ما ستحصل عليه من جوجل فهو لا يدعم البحث الفائق wild card بمعـــنى أنه لا يعطيك كل مترادفات الكلمة أو مشتقاتها، وإنما يعطيك فقط ما أردت البحث عنه بالتمام والكمال .

كذلــك يســتبعد هذا المحرك كل الكلمات التى قد لا يكون هناك معنى لها، مــثل "و" أو " com" أو "عند"، بمعنى آخر يستبعد تلك الكلمات التى لا تساعد فى تضييق نطاق البحث، وإنما قد تؤخر نتيجة البحث أو تعمل على اتساعه بدرجة قد تــبدو معــه النــتائج باهتة و لا تمت للكلمات المبحوث عنها بصلة.ومن هنا تجدر الإشــارة إلى أهمية أن تكون الكلمات التى يتم البحث عنها مصاغة بدقة عالية وأن لا تحتوى كلمات ليست ذات معنى أو أنها يمكن أن تقتل نتيجة البحث.

يقبل جوجل استخدام بعض علامات التمييز فى اللغة مثل رموز الإضافة + والناقص – والشارحة :، إضافة إلى استخدام بعض الحروف التى تعنى شيئا محددا كالتالي:

1- استخدم علامة الإضافة + لتعنى البحث عن عبارة محددة تريدها أن تظهر فى نتيجة البحث كما هى

2- استخدم علامة الناقص(الطرح) – لتعنى بها استبعاد بعض الكلمات التى لا تريدها أن تظهر فى نتيجة البحث

3- استخدم علامة الشارحة: لتعنى بها البحث عن موقع محدد ليأتي بعدها اسم الموقع الذي تريده، وتكتب هكذا site:sampledomain.com

4- اســتخدم علامة التنصيص "" لتعنى بها البحث عن عبارة محددة تريدها أن تأتى كما كتبتها، ويقول مؤسسو جوجل أن "هذه الطريقة تنفع خصوصا عند البحث عن أقوال شهيرة أو أسماء محددة".

5- للبحــث عــن معلومــات صــحفية يمكــنك أن تكتــب الــتالى press site:www.google.com ثــم تضــغط على مربع Goويعنى ذلك أنك تبحث عن معلومات صحفية داخل موقع جوجل.

السمات العامة للمحرك:

يتمــيز جوجــل بقدرته الهائلة على البحث فى لغات متعددة، والنتائج التى يخــرج بها غالبا ما ترضى الباحث، إضافة إلى بعض المميزات الأخرى المتعلقة بإمكانية البحث فى الصور ومجموعة من الملفات النصية مثل ملفات PDF، وربما يفيد هنا أن نذكر ما يشير إليه القائمين على هذا المحرك عند البحث بالصور.

"البحث عن الصور في Google هو الأشمل في الويب، إذ يوجد أكثر من 250 مليون صورة مفهرسة وجاهزة للعرض. لكي تستعمل البحث عن الصور، اذهب إلى صفحة "بحث متقدم" أو إلى http://www.google.com/imghp?hl=ar وأدخل عبارة بحث في مربع البحث عن الصور، ثم انقر زر "ابحث". وفي صفحة النتائج، يكفي أن تنقر على الصورة المصغرة كي تراها أكبر، وكذلك صفحة الويب حيث توجد الصورة. (ملاحظة: يجب أن تُجعل الإنجليزية لغة واجهة الاستخدام في صفحة "تفضيلات").

قد تكون الصور التي تجدها عبر خدمة "البحث عن الصور" في Google خاضعة للحماية بموجب حقوق الطبع والنشر. ومع أنه بإمكانك الوصول إلى هذه الصور عن طريق هذه الخدمة التي نقدمها، لا يمكننا أن نمنحك أية حقوق في استعمالها لأي غرض غير مشاهدتها في الويب. لذلك إذا أردت أن تستعمل أية صور وجدتها بواسطة خدمتنا هذه، ننصحك بالاتصال بصاحب الموقع لتحصل على الإذن منه".

إن هذه العبارة داخل جوجل تعنى الكثير فيما يتعلق بالبحث بالصور وحقوق الملكية الفكرية لهذه الصور وأهمية الحصول على إذن صاحب الصورة

قبل التصرف فيها سواء بنشرها أو استخدامها فى أغراض أخرى، والحذر عند البحث بها من قبل الأطفال، وأهمية أن يراعى الكبار ذلك دائما عند إعطاء صغارهم الإذن باستخدام محركات البحث عند البحث عما يرغبون به.

أيضا من مميزات جوجل قيامه بالترجمة لبعض نتائج المواقع، ويمكنك عزيزي القارئ أن تكتشف قدرات جوجل فى هذا المجال وتحكم عليها بنفسك.

كما يتمكن جوجل أيضا من التعرف على الصفحات المختلفة المكتوب بها النص، فإذا كان مكتوبا ببرنامج Acrobat PDF فسوف يمكنك من التعرف غلى محتوياته إذا كان هذا البرنامج محملا على جهازك أو ينقلك إلى موقع Acrobat ليمكنك من تحميله ومن ثم استخدامه وتظهر علامة PDF بجانب الصفحات التى يمكن أن يكون مكتوبا بها فى نتائج الموقع التى تبدو أمامك.

ربما ينبغي الإشارة أيضا إلى أن جوجل يقوم بعرض صفحات مخبأة عن المواقع التى قد تظهر فى نتيجة البحث ولا يمكنك الولوج إليها، هنا عليك استخدام الواصلة "نسخة مخبأة" حيث يقوم جوجل بالاحتفاظ بصورة من الصفحة فى مخزنه

الإلكتروني ليمكنك من مشاهدة هذه الصفحة إذا تعذر عليك رؤيتها لسبب أو لآخر لو اتبعت الواصلة الطبيعية التى أتاحها المحرك أمامك.

كما أنه يستخدم أيضا مصطلح "ضربة حظ" ليمكنك من إجراء بحث سريع للغاية يحمل نتائج سريعة قد تحتاجها لأن الوقت أمامك محدود أو لسبب آخر.

نتائج البحث فى المحرك:

عندما تستخدم البحث بجوجل عليك أن تلاحظ دائما أنه يعطيك نتيجة البحث محـــددا الفئة التى قمت فيها بالبحث وعدد الصفحات التى عثر عليها، وأيضا زمن البحث المستغرق.

كما أنه مع كل نتيجة بحث يمنحك المعلومات التالية:

1- الواصلة أو الارتباط للصفحة التى ظهرت كنتيجة (LINK)

2- وصف للصفحة

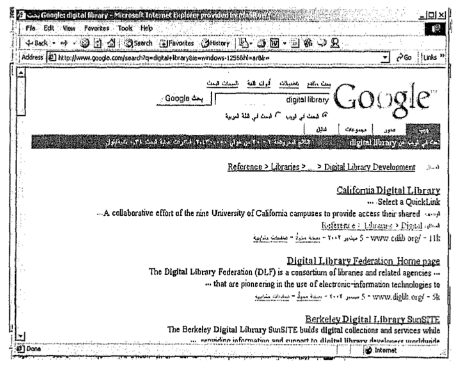
3- مجـــال الصـــفحة (المجال الموضوعي الذي تنتمي إليه الصفحة والتي قام جوجل بتصنيفها موضوعيا فيه)

4- حجم الصفحة

5- تاريخ الصفحة

6- النسخة المخبأة

7- الصفحات المشابهة (من الناحية الموضوعية)



طرق البحث المستخدمة فى جوجل :

لا تكاد محركات البحث تختلف فيما بينها إلا فى فروقات بسيطة فى أساليب البحث المستخدمة، فجوجل يدعم أيضا استخدام اللغة الطبيعية والبحث بجمل واستخدام أساليب المنطق البوليني، وإن كان مظهر استخدام هذا الأسلوب يختلف من محرك إلى آخر مثل استخدام كلمة and أو التعبير عنها برمز الإضافة +،

ولكـــن عمومــــا هناك شبه اتفاق فيما يتعلق باستخدام الأساليب المختلفة للبحث بين المحركات على الإنترنت.

ينبغي الإشارة فى النهاية إلى أنه يمكنك البحث ضمن نتائجك التى ظهرت مـــن خــــلال إضافة كلمة أخرى إلى البحث نفسه فيقوم جوجل بتنقية نتائج البحث بحيث تحتوى على الكلمة التى وضعتها إضافة إلى الكلمات الأخرى، وينبغي دائما الاهــتمام بهـــذا الـــنوع من النتائج لأنه يقترب من متطلبات البحث فى المحرك، والتـــنويه على أنه إذا أراد القيام بالبحث مرة أخرى فعليه إما أن يقوم بالبحث فى نتائج البحث الذي قام به من قبل أو أن يقوم ببحث جديد تماما.

تتبقى بعض الإشارات التى ينبغي التنويه إليها وهى تتعلق بالأسئلة الشائعة للباحثيـــن حـــول جوجل وإمكاناته ويمكنك عزيزي القارئ أن تجدها فى الصفحة التالية داخل جوجل http://www.google.com/intl/ar/faq.html وذلك للقراء باللغة العربية.

الفصـــل الـــرابـــع

محرك بحث

" Alltheweb "

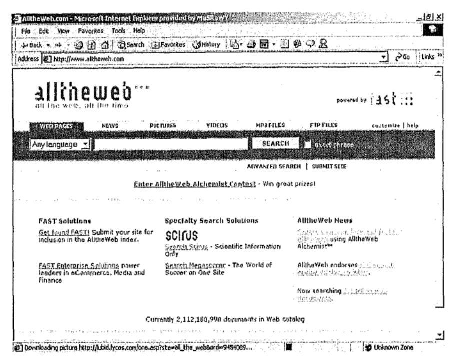
الفصل الرابع

محرك بحث " *Alltheweb* "

- محرك البحث *Alltheweb*:

يسمى هذا المحرك باسم "أول ذى ويب" أو بالعربية "كل ما على الصفحات العنكبوتية"، كما أنه أحيانا يحمل اسم الشركة التى قامت بتصميمه وتنفيذه وإطلاقه وهى شركة فاست (FAST) Fast Search & Transfer ASA

يمتاز هذا المحرك بأنه يمكن البحث فيه باستخدام مجموعة من الخيارات، ويمكن البحث بلغة معينة (يضع 49 لغة يمكن البحث بأي منها، أو بها جميعا)، كما أنه يبحث فى حوالي 2.1 بليون صفحة، ويقول مؤسسوه أنه يحتوى على صفحات أكثر من جوجل يمكن البحث فيها، حيث تصل عدد صفحات جوجل إلى 2.073418204 بينما تصل عدد صفحاته هو إلى 2.095568809، بينما تجرى فيه عملية التحديث دوريا خلال فترة تتراوح من 7 إلى 11 يوما.



شـــكل صـــفحة محرك alltheweb "كل الصفحات العنكبوتية" العربية على شبكة الإنترنت

http://www.alltheweb.com/

خصائص المحرك:

يوفر هذا المحرك فرصة تحديد للبحث عن صفحات الويب، أو الصور، أو الفـــيديو، أو ملفـــات الصوت MP3، تحميل برمجيات معينة عبر بروتوكول نقل الملفات FTP، بالإضافة إلى أنه يوفر مجموعة قوية من طرق البحث منها البحث بكلمة مفردة أو جملة معينة، أو البحث في عنوان محدد، أو البحث في موقع معين، أو عـــن وصـــلة معيـــنة داخل أحد المواقع. ويمكن لمستخدمي هذا المحرك تحديد

طـــريقة البحـــث المناسبة لاحتياجاتهم.وهو يضم مجموعة من الفئات يمكن البحث عنها هى:

1- صفحات الشبكة العنكبوتية Web pages

2- الأخبار News

3- الصور Pictures

4- أفلام الفيديو Video

5- الملفات الصوتية من نوع MP3

6- ملفات بروتوكول نقل الملفات FTP Files

ويـــتم البحث هنا مباشرة على المجموعة الأولى من الخيارات الستة وهى صـــفحات الشـــبكة العنكبوتـــية، إلا إذا أردت البحـــث فـــى مجموعة أخرى من المجموعات الست الباقية.

ربمـــا يتفرد هذا المحرك هنا بنوعين من البحث هما استخدام البحث على المجموعـــة الخامســـة " الملفـــات الصـــوتية مـــن نوع MP3" والسادسة " ملفات بـــروتوكول نقل الملفات FTP Files "، إضافة إلى قدراته فى البحث عن الصور

وأفـــلام الفيديو، وهو ما يمكن أن نشير إليه هنا بأنه محرك يبحث فى جميع أنواع أوعـــية ومصـــادر المعلومـــات ســـواء المقروءة أو المسموعة أو المرئية، الثابتة والمتحركة.

خصائص البحث فى المحرك:

ســـبقت الإشـــارة إلى أن هذا المحرك ينفرد بمجموعة من القدرات اللغوية تبدو فى قدرته على التعامل بعدد 49 لغة فى آن واحد أو بشكل منفصل، وهو ما قد لا يـــتوافر فـــى محـــركات بحث أخرى على شبكة الإنترنت، وإن كانت قدراته لا تقارن هنا بقدرات جوجل اللغوية الذي يتعامل مع 79 لغة.

يدعـــم محـــرك "فاســـت أو كل الإنترنت" البحث البسيط والبحث المركب، بمعنى أنه يمكنك أن تكتب ما تريد فى مربع البحث Search Frame وسوف يقوم المحرك بتنفيذ البحث الخاص بك.

يتميز المحرك بقدراته القوية فيما يعرف بالبحث المتقدم، فهو يقدم مجموعة الخيارات التالية:

1- البحث باللـــغة أو بشفرة لغويـــة محددة كالشفرة الموحـــدة (Unicode)

2- يستخدم المعاملات المنطقية + أو – كالتالي :

- يجب أن تحتوى الجملة على التاليmust include
- من المحتم أن تحتوى الجملة على التالى Should include
- لا يجب أن تحتوى الجملة على الكلمة التالية Should not include

ســواء وردت هــذه الكلمة فى عنوان الصفحة أو فى عنوان الموقع أو فى النص الكامل للصفحة أو فى عنوان الموقع المضيف أو فى عنوان الوصلة لعنوان الموقع.

إضافة إلى أن الموقع الذي يتم البحث عنه قد يحتوى على ملفات (صور أو نصــوص) مدمجــة فيه كملفات الصور و الصوت والموسيقى والرسوم المتحركة والملفات النصية الأخرى من نوع pdf مثلا.

أيضــا يمكــن تحديد نوع النطاق Domain ما إذا كان تجارى com أو تعليمــي edu أو غيره. وأيضا القارة أو الدولة التى ينتمي إليها هذا النطاق وأيضا يمكن البحث باستخدام عنوان IP Addres.

كمـــا يمكـــن تحديد نوع الشكل والقالب الذي يمكن أن تكون فيه المواقع أو الملفـــات المطلوبـــة، وتواريخ تحديثها، وحجم الملف المطلوب سواء كان أكبر من حجم معين أو أصغر .

أيضـــا مـــن الإمكانات الخلابة فى هذا المحرك أنه يمكنك أن تقوم بتحديد عمق الصفحة الذي تريدها داخل الموقع، وإلى أي مستوى يمكن للمحرك أن يبحث، فمـــن المعـــروف أن صفحات الوب تبنى بشكل هرمي ويتم الربط بينها من خلال علاقـــات محددة هى عناوين الصفحات أو الوصلات التى تقع فيها، وهذا المستوى الهرمـــي من البناء يمكن من وضع مستوى محدد لكل مجموعة من الصفحات على المستويين الأفقي والرأسي.

أيضـــا يحـــدد البحث المتقدم إمكانات البحث فى الصفحات الشخصية فقط Personal Homepages

السمات العامة للمحرك:

ســـبقت الإشارة إلى القدرات اللغوية لهذا المحرك، إضافة أيضا إلى أنه يفرد مجموعـــة مـــن إطـــارات البحث لكي تقوم بالبحث عن مصادر معينة على شبكة الإنترنـــت كملفـــات الصـــور والفـــيديو وملفات الصوت كملفات mp3 الصوتية،

والملفـــات الإخباريــــة، وملفات بــــروتوكول نقل الملفات FTP (File Transfer Protocol).

نتائج البحث فى المحرك:

عندما تستخدم البحث بفاست عليك أن تلاحظ دائما أنه يعطيك نتيجة البحث مبينا فيها عدد الصفحات التى تحتوى على نفس المصطلح الذي قمت بالبحث عنه، بادئا بعرض أي أخبار قد تم تكشيفها خلال الساعات الماضية تتعلق بالموضوع، ثم يبدأ بعرض نتائج البحث كالتالي:

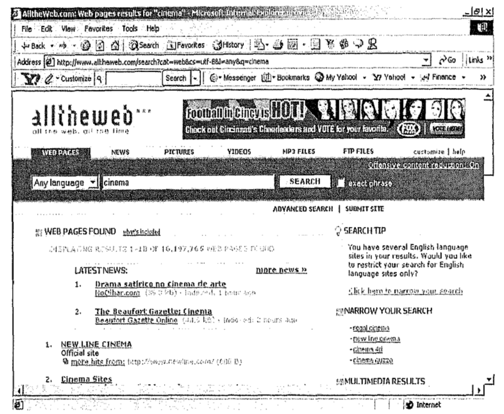
1- الواصلة أو الارتباط للصفحة التى ظهرت كنتيجة (LINK)

2- مجتزأ من الصفحة (أول نص فيها على سبيل المثال)

3- وصف للصفحة

4- مجـــال الصـــفحة (المجال الموضوعي الذي تنتمي إليه الصفحة والتي قام فاست بتصنيفها موضوعيا فيه)

5- حجم الصفحة



http://www.alltheweb.com/

طرق البحث المستخدمة في فاست :

لـــيس هناك ما يمكن إضافته إلى ما سبق الإشارة إليه، ويمكن الرجوع إلى الموقع نفسه للحصول على مزيد من الإجابات عن طرق البحث المستخدمة، وإن كـــان يجـــب أن أشـــير إلى أن هذه المحركات تفاجئنا كل يوم بطرق بحث جديدة، وبالتالي فليس هناك حقيقة نهائية مؤكدة حول طرق البحث.

الفصـــل الـــخـــامـــس

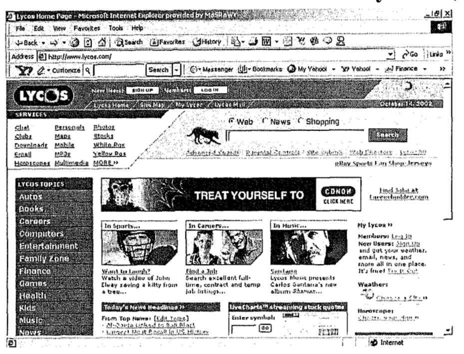
محرك بحث

" Lycos "

الفصل الخامس

محرك بحث " *Lycos* "

- محرك البحث Lycos :



http://www.lycos.com

تم إطلاق هذا المحرك للمرة الأولى فى مايو عام 1994، وقد قام بتطويره كل من د. مايكل مولدين ثم باعه لشركة أمريكا أونلاين America Online، وفى عـام 1995 أصـبحت لايكوس شركة مساهمة يساهم فيها كل من شركة CMG وجامعـة كارنيجى ميلون، وفى عام 1996 تم وضعها فى البورصة، وفى فبراير 1998 قامـت شركة الشبكات الأمريكية USA Networks بشراء حوالي 61%

من أسهم لايكوس بقيمة تبلغ 18 بليون دولار، وأصبحت هذه الشركة تعرف الآن باسم USA/Lycos Interactive Network.

يحتوى كشاف هذا المحرك على 38 مليون موقعٌ تقريبا، وهو لا يدعم تركيبات Tag Meta، أو الإطارات، أو خرائط الصور، أو التعليقات داخل صفحة HTML، وهو يوفر مجموعة من وسائل البحث تتفق مع المحركات السابق ذكرها.

ويمكن استخدام وسيلتين أيضا عند البحث في لايكوس:

- البحث مباشرة في قائمة مكونة من أكثر من 21 موضوعا.
- استخدام محرك البحث (أو إطار البحث) حيث يوفر فرصة البحث في كل الوثائق، أو في مجموعة محددة من المواقع، أو في عناوين المواقع فقط.

وتضم خيارات البحث في أي نوع من المحتويات: مثل البحث عن عناوين، أو مؤلفو كتب معينة مع إمكانية شرائها عبر شبكة الإنترنت، أو البحث بهدف تحميل برمجيات، أو البحث في الموسيقى (المسموعة)، أو البحث في الأخبار، أو

البحـــــث في الصفحـــات الشخصية Personal home pages وكذلك البحث في الصور و الطقس والمؤشرات الاقتصادية...الخ.

خصائص المحرك:

يوفـــر هـــذا المحـــرك فرصة تحديد للبحث عن صفحات الويب، والأخبار وغيرها، ولكنه بصفة أساسية يحدد البحث فيه فى المجالات الثلاثة التالية:

1- صفحات الشبكة العنكبوتية Web pages

2- الأخبار News

3- التسوق Shopping

خصائص البحث فى المحرك:

يـــتعامل هـــذا المحـــرك مـــع 85 لغة منها العربية وبذلك فهو يعد أقوى المحركات المتاحة بلغات مختلفة، وهو يستعمل نفس أسلوب محرك " كل صفحات الإنترنت" فهو فى البحث المتقدم يمكنك من البحث بالأساليب الثلاثة التالية:

a. يجب أن تحتوى الجملة على التاليmust include

b. من المحتم أن تحتوى الجملة على التالى Should include

c. لا يجب أن تحتوى الجملة على الكلمة التالية Should not include

ســـواء وردت هـــذه الكلمة فى نص الصفحة، أو فى عنوان الصفحة أو فى عنوان الموقع أو فى وصلة لعنوان الموقع.

كمـــا أنه يمكن أيضا من استبعاد مواقع معينة حيث يمكنك ذكر اسم الموقع ونطاقه، أو تضمين مواقع معينة يمكنك البحث فيها مباشرة. .

كمـــا أنه يوفر وسيلة لتنقيح نتائج البحث باستبعاد المواقع التى تحتوى على صـــفحات خارجة بدرجات محددة من التنقيح، كأن يقوم المحرك بالاستبعاد الكلى لهذه المواقع، أو تركها كما هى، أو اختيار بعضها والاحتفاظ به فى نتائج البحث. .

السمات العامة للمحرك:

يتمتع المحرك بميزتين أساسيتين:

1- البحث فى مواقع محددة أو استبعاد مواقع غير مرغوب البحث فيها.

2- تنقيح الصفحات التى قد تحتوى على نتائج تعتبر خارجة عن إطار العرف والتقاليد والدين.

لكنه يعتبر محركا تجاريا من الأساس لأنه يوفر وسيلة أساسية فيه للتسوق عبر الإنترنت.

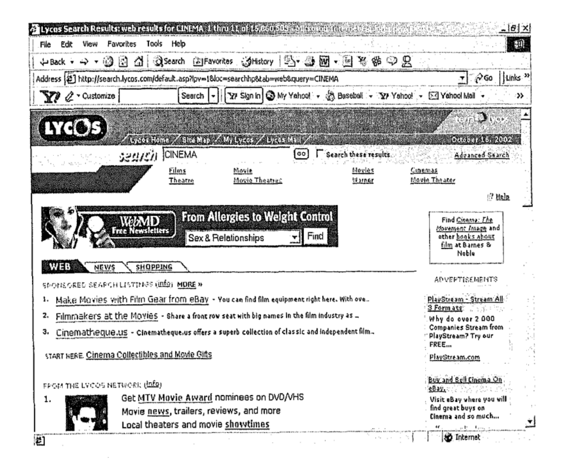
نتائج البحث فى المحرك:

عـــندما تســـتخدم البحـــث لايكوس عليك أن تلاحظ دائما أنه يعطيك نتيجة البحث مبينا فيها عدد الصفحات التى تحتوى على نفس المصطلح الذي قمت بالبحث عـــنه، بادئـــا بعـــرض أي أخـــبار قـــد تم تكشيفها خلال الساعات الماضية تتعلق بالموضوع، ثم يبدأ بعرض نتائج البحث كالتالي:

1- الواصلة أو الارتباط للصفحة التى ظهرت كنتيجة (link)

2- مجتزأ من الصفحة (أول نص فيها على سبيل المثال)

3- عنوان الصفحة address



http://www.lycos.com/

الفصل السادس

محرك بحث

" Excite "

الفصل السادس

محرك بحث " *Excite* "

- **محرك البحث إكزايت Excite :**

يحتوي كشاف محرك البحث Excite excite. على 30 مليون موقع تقريبا، وغالبا ما يبحث فى المواقع التى تحتوى على صفحة واحدة، وهو لا يدعم تركيبات Tag Meta أو الإطارات أو خرائط الصور، أو التعليقات داخل صفحة HTML. كما أنه يقدم مجموعة جيدة من أدوات البحث والتجول عبر الإنترنت. وتحتوي قائمة البحث على 18 رأس موضوع يمكن البحث من خلالها مباشرة. ويمكن القول بأن البحث في مثل هذه القوائم عادة ما يكون مقيدا بنتائج محددة سلفا. ويقول مؤسسو محرك Excite على موقعهم على الإنترنت " بأنه محرك بحث مبني بطريقة منهجية ثورية للبحث على الإنترنت، والحصول على نتائج يتوقعها مستخدمو الإنترنت".



http://www.excite.com

خصائص المحرك:

يقـــوم محرك إكزايت بالبحث فى أربعة أنواع من مصادر المعلومات على شبكة الإنترنت هى:

1- صفحات الشبكة العنكبوتية Web pages

2- الصور Photos

3- دليل إكزايت Directory

4- الأخبار News

خصائص البحث فى المحرك:

قد لا تبدو أي هناك أي مشاكل للكثير منا الآن فى البحث عن الأخبار على محـــركات البحث، لكن يظل لكل محرك إمكاناته التى ينفرد بها، ففي هذا المحرك يمكـــنك كـــتابة الموضوع الإخباري الذي تريده، على أن تختار أسلوب البحث فى المعلومات الإخبارية.

أما بالنسبة للبحث فى الصور فهو لا يختلف كثيرا عن البحث فى الأخبار، يكفـــى كـــتابة الموضوع المطلوب البحث عن صور له، ويسبق ذلك تحديد أسلوب البحث فى الصور.

وبالنسبة للبحث فى دليل إكزايت، فيجب أن تكون على معرفة كافية وتامة بما تقوم بالبحث عنه وذلك فى القائمة الشجرية للموضوعات Directory وهو دليل بمجموعـــة من الموضوعات المختارة من قبل العاملين فى المحرك يمكن الاختيار مـــنها والتنقل من موضوع آخر فيها بشكل هرمي، بمعنى أن تقوم بالبحث فى كل الموضـــوعات المرتـــبطة بالموضـــوع الرئيســـي الذي ترى أنه قد يحتوى على الموضوع الذي تريد البحث عنه، وقد يوضح ذلك المثال التالي:

1. Sports > Football
2. Sports > Football > Arena
3. Sports > Football > Soccer
4. Sports > Gaelic > Football
5. Sports > Fantasy > Football
6. Sports > Football > American
7. Sports > Football > Canadian
8. Shopping > Sports > Football

فنظرة واحدة على هذه الطرق من البحث يمكن أن تدلنا جميعا على المقصود من هذه الوسيلة فى البحث إلا إذا كنت تبحث عن كرة القدم كرياضة فيمكنك أن تبحث عنها تحت الموضوع الرئيسي (رياضة)، وإذا كنت تبحث شراء كرة قدم فيمكن أن تبحث عنها فى الموضوع الرئيسي (التسوق)، وإذا كنت تبحث عن الحوادث الرياضية الغربية فى مجال كرة القدم فإن النموذج الخامس يبين لنا ذلك (رقم 5 فى القائمة السابقة) وهكذا.

يستخدم اكزايت أيضا المعامل المنطقي البولياني "و" and، ويمكنك استخدامها فى شكلها اللغوي كحروف أو استخدامها على شكلها الرمزي كعلامة +،

وكذلك استخدام حرف أو or للتخيير بين موضوعين أو مصطلحين كذلك يستخدم اكزايت المعامل البوليانى ليس Not مثل

الريحان ليس الياسمين

بمعنى أنك تستبعد الموضوع الثاني ولا تريد ظهوره فى نتائج البحث، بينما تريد لنتائج البحث كلها أن تحتوى على الريحان فقط.

السمات العامة للمحرك:

يتمتع هذا المحرك بمجموعة من الميزات، مثل البحث فى الصور والأخبار، لكنه يقدم مجموعة كبيرة من الخدمات الأخرى حيث يعمل كبوابة Portal ويقدم مجموعة كبيرة من الخدمات الاقتصادية والاجتماعية يمكنك ملاحظتها بنفسك.

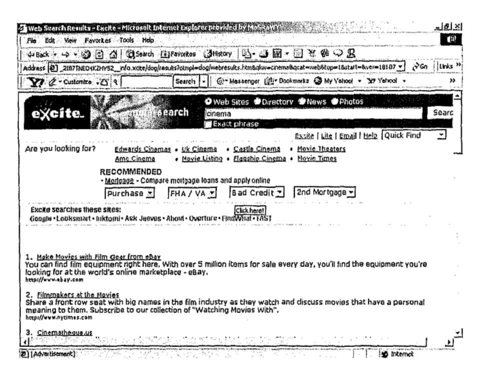
نتائج البحث فى المحرك:

يوفر إكزايت أمامك ثلاث قطع من المعلومات ففي كل نتيجة لموقع يتفق مع ما قمت بالبحث عنه :

1. الواصلة أو الارتباط للصفحة التى ظهرت كنتيجة (link)

2. مجتزأ من الصفحة (أول نص فيها على سبيل المثال)

3. عنوان الصفحة address



http://msxml.excite.com

وهــو يقــوم بترقيم النتائج فى نفس الوقت، لكنه لا يوفر أي معلومات عن العدد الإجمالي للنتائج الذي ظهر.

الفصـــل الســـابع

محرك بحث

" Altavista "

الفصل السابع

محرك بحث " *Altavista* "

- محرك البحث Altavista :

خصائص البحث:

يعد ألطافيستا من أشهر محركات البحث في العالم لمستخدمي شبكة الإنترنت حيث أنه يوفر مواقع ومعلومات أكثر اتصالا بموضوع البحث – علي حد وصف الموقع – بسرعة وكفاءة، وينعكس ذلك على نتائج البحث التي تتراوح ما بين صفحات الوب، وملفات وسائط متعددة(صور / صوت/ مقاطع فيديو)، مواقع تسوق، أخبار، بالإضافة إلى أنه يوفر خدمة ترجمة مجانية لنتائج البحث بأكثر من لغة. ويتميز محرك بحث altavista بتكنولوجيا عالية الأداء لخدمات البحث في النص الكامل.



http://www.altavista.com/

خصائص البحث فى المحرك:

يمكن استخدام وسيلتين للبحث في هذا المحرك هما:

1. محـــرك البحـــث الذي يوفر طريقة للبحث في لغات معينة على الشبكة العنكبوتية لأكثر من 25 لغة منها الصينية والسويدية وغيرهما. ويتميز هـــذا المحرك بإمكانية البحث المباشر بجمل طبيعية، بمعنى عدم أهمية إســـقاط حـــروف الوقـــف وغيرها، كما أنه يعمل بأساليب البتر السابقة واللاحقة
2. مـــن خلال قائمة البحث الموضوعية أو المجالات الرئيسية للبحث والتي تبلغ **14** مجالا رئيسيا.

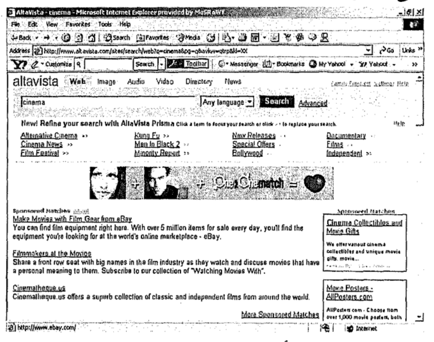
كمـــا تـــتوافر خدمـــات إخبارية مباشرة من خلال العناوين الإخبارية التي يوفرها مباشرة أمام المستخدم.

السمات العامة للمحرك:

ويمكن الإشارة أيضا بأن الشركة المالكة لهذا المحرك تشكو من قيام بعض المســـتخدمين بوضـــع واصفات فى صفحاتهم لا تعبر عن ما يحتويه مواقعهم، مما يؤثـــر علـــى عملية الاسترجاع، ويقود إلى ظهور نتائج لا تتفق مع استراتيجيات

البحـــث التى تم استخدامها[1]، أو قيام بعض المواقع بمضاعفة عدد نفس الصفحات فى الموقع، أو قيام بعض المواقع بإنتاج بعض الصفحات بشكل آلي على الرغم من عـدم وجـــود أي محـــتويات بهـــذه الصفحات، مما يدفع المستخدم للنقر فى هذه الصفحات للولوج لصفحات أخرى، أو وجود صفحات الهدف منها فقط وضع بعض الروابط التى تدفع المستخدم للنقر فيها للولوج لصفحات أخرى. ويعتبر هذا المحرك ثاني أكبر محرك على الإنترنت.

نتائج البحث فى المحرك:



شكل صفحة نتيجة البحث على ألطافيستا

http://www.altavista.com/

---

[1] Search Engine Tutorial for Web Designers: AltaVista. url <http://northernwebs.com/set/alta_vista.html>

يعـــرض هذا المحرك نتائج البحث بشكل بسيط، لكنه يضيف مع كل موقع يعرضه كنتيجة عنصرين، هما الصفحات المتشابهة مع الصفحة التى تظهر كنتيجة أو ذات العلاقة بها، وإمكانية الترجمة المباشرة للصفحة التى تظهر كنتيجة.

بانتهائـــنا مـــن استعراض هذه المحركات الخمس يمكننا أن ننتقل الآن إلى تجربتنا المتعلقة بالمقارنة معمليا بين هذه المحركات.

قياس أداء المحركات المختارة :

تم إجراء بعض البحوث من خلال استخدام إطار البحث Query Box في المحركات الستة المختارة لقياس مجموعة من الاحتمالات التي يمكن أن تجيب على الأســئلة والملاحظـــات التـــي أثرناها أثناء مناقشتنا لقدرات كل محرك حيث قمنا بإلقـــاء ثلاثـــة أســئلة على كل محرك بحث، الأول باستخدام كلمة مفردة، والثاني باستخدام تركيب بولياني، والثالث باستخدام اللغة الطبيعية، وهذه المصطلحات هى كالتالي:

1- المصطلح الأول: Library (مكتبة)

2- المصطلح الثاني: Digital Library(مكتبة رقمية)

3- المصـــطلح الثالث: Digital Reference collection (مجموعة مراجع رقمية)

وذلك بهدف قياس، وتقييم مجموعة من الخصائص لكل محرك بحث على حـدى، ثـم مقارنتها ببعضها البعض. إضافة إلى استخدام ساعة ميقاتية لإحصاء المدى الزمني بين تلقي الاستفسار والإجابة عليه. وكانت النتائج كالتالي :

جدول(1) عدد المواقع والصفحات وزمن البحث لثلاثة أبحاث على المحركات الستة المختارة

المحرك	المصطلح	Library	Digital Library	Digital Reference Collection	المجموع	المتوسط
Northern light	العدد	13.948.273	1.191.725	305.401	15.445.399	5.148466
	الزمن	6 ث	17 ث	9 ث	32	10.7ث
Google	العدد	35.800.000	1.400.000	785.000	37.985.000	12.661.600
	الزمن المرصود الزمن الذى حدده الموقع	14 ث 0.66 ث	2. 0 ث 0.20 ث	7.0 ث 0.72	23ث 1.58	7.8ث 0.52
Altavista	العدد	50.278.930	19.722.910	58.974.970	128.976810	42.992270
	الزمن	6 ث	6 ث	7 ث	19ث	6.33ث
Excite	العدد	8.605.302	4.173.785	4.726.105	17.505192	5.835064
	الزمن	8 ث	5 ث	16 ث	29ث	9.6ث
Lycos	العدد	16.359.892	1.285.139	256.497	2.020.528	673.509
	الزمن	4 ث	6 ث	3 ث	13ث	4.3ث
All the web	العدد	15.179.544	1.164.324	257.428	16.601296	5.447956
	الزمن	3 ث	3 ث	2 ث	8ث	2.7ث

بــناء علـــى هذا الجدول يتضح أن محرك *All The Web* هو صاحب أقل زمـن بحثي بين المحركات الستة، بينما يحتل *Altavista* المرتبة الأولى في عدد الصفحات الناتجة عن البحث. ويمكن الإشارة إلى أن *Google* يضع زمن البحث فـي مقدمة نتائج البحث الذي قام به، والاختلاف بين الزمنيين فى البحث ناتج عن سـرعات الــنقل للبيــانات، وسرعة المعالج، وزمن التحميل، إضافة إلى أن زمن البحث الذي يذكره هذا المحرك هو زمن البحث فى قاعدة بياناته.

جدول (2) قياس الواقعات العشر الأولى لكل محرك وفقا لدولة المنشأ / أو لجنسية الموقع

(البحث عن كلمة مفردة) Library

م	المحرك / الدول التي تنتمي إليها المواقع	أمريكا	بريطانيا	استراليا	المجموع
1	Northern light	10	-	-	10
2	Google	7	1	-	8
3	Excite	10	-	-	10
4	AltaVista	8	1	1	10
5	All the web	9	1	-	10
6	Lycos	10	-	-	10
	المجموع	54	3	1	58

يلاحــظ مــن جدول(2) أيضا سيطرة المواقع الأمريكية بنسبة 93.1% ثم بريطانيا بنسبة 5.1% ثم استراليا بنسبة 1.8%

جدول (3) قياس الواقعات العشر الأولى لكل محرك وفقا لنوع الموقع

(البحث عن كلمة مفردة Library)

م	المحرك	namewww.site					مجموع
		Edu	com	gov	org	أخرى	
1	AltaVista	2	7	1	–	–	10
2	Lycos	3	5	1	1	–	10
3	Excite	–	8	1	–	1	10
4	All the web	3	4	1	2	–	10
5	Google	5	2	1	2	–	10
6	Northernlight	2	1	2	2	3	10
	المجموع	15	27	7	7	4	60

يلاحـــظ بناء على جدول (3) أن مواقع الشركات تسيطر بنسبة 45% من إجمالي عدد المواقع عند البحث بكلمة واحدة مفردة، تليها المواقع التعليمية التابعة للجامعات والكلـــيات والمـــدارس بنسبة 25% ثم المواقع الحكومية، ومواقع المنظمات بنسب متساوية 11.6%.

جدول (4) قياس الواقعات العشر الأولى لكل محرك وفقا لدولة المنشأ / أو لجنسية الموقع

Digital library(البحث عن كلمتين)

م	المحرك / الدولة	أمريكا	إنجلترا	ألمانيا	استراليا	كندا	نيوزيلندا	مجموع
1	AltaVista	10	–	–	–	–	–	10
2	Lycos	10	–	–	–	–	–	10
3	Excite	10	–	–	–	–	–	10
4	All the web	10	–	–	–	–	–	10
5	Google	9	–	–	–	–	1	10
6	Northernlight	9	–	1	–	–	–	10
	المجموع	58	–	1	–	–	1	60

فـــى الجـــدول (4) تحـــتل أمريكا أيضا نفس العدد من المواقع مثلما هو الحال فى الجـــدول (2) وإن كانت كل من ألمانيا ونيوزيلندا تظهران فى الصورة ولكن بشكل متواضع .

جدول (5) قياس الواقعات العشر الأولى لكل محرك وفقا لنوع الموقع
(البحث عن كلمتين) Digital Library

م	المحرك / نوع الموقع	namewww.site					مجموع
		Edu	com	gov	org	أخرى (NET&DOT)	
1	Northern light	7	–	1	1	1	10
2	Google	2	–	3	3	-	8
3	Excite	2	2	1	4	1	10
4	AltaVista	1	2	6	-	1	10
5	All the web	–	1	2	6	1	10
6	Lycos	2	4	2	2	-	10
	المجموع	14	9	15	16	4	58

يلاحظ من جدول (5) تواجد مواقع المنظمات (.org) أولا وبنسبة 27.6% ثم المواقع الحكومية (.gov) بنسبة 25.9% ثم المواقع التعليمية (.edu) بنسبة 24.1% ثم مواقع الشركات بنسبة 15.5%

جدول (6) قياس الواقعات العشر الأولى لكل محرك وفقا لدولة المنشأ / أو جنسية الموقع

Digital Reference Collection (البحث عن جملة)

م	المحرك / الدولة	أمريكا	هولندا	إنجلترا	فرنسا	استراليا	نيوزيلندا	كندا	كرواتيا	المجموع
1	AltaVista	10	–	–	–	–	–	–	–	10
2	Lycos	9	–	–	–	1	–	–	–	10
3	Excite	9	–	1	–	–	–	–	–	10
4	All the web	9	–	–	–	1	–	–	–	10
5	Google	10	–	–	–	–	–	–	–	10
6	Northernlight	4	1	–	1	1	2	–	1	10
	المجموع	51	1	1	1	3	2	–	1	60

يلاحظ أيضا من واقع هذا الجدول (6)سيطرة مواقع الولايات المتحدة الأمريكية بنسبة ساحقة تصل إلى 85% من إجمالي عدد المواقع التى عرضها البحث تليها استراليا بنسبة 5% ثم نيوزيلندا.

جدول (7) قياس الواقعات العشر الأولى لكل محرك وفقا لنوع الموقع

Digital library Collection(البحث عن جملة)

م	المحرك / الصفة	namewww.site					مجموع
		Edu	com	gov	org	أخرى	
1	AltaVista	2	3	-	5	-	10
2	Lycos	7	-	-	3	-	10
3	Excite	2	1	-	7	-	10
4	All the web	2	-	2	6	-	10
5	Google	4	-	-	5	1	10
6	Northernlight	3	3	-	2	2	10
	المجموع	20	7	2	28	3	60

فيما يتعلق بالبحث عن جملة، فقد كانت نتائج البحث بالنسبة لأنواع المواقع (وفقا لجدول 7)ظهور مواقع المنظمات الدولية والإقليمية وغيرها من المنظمات غير الحكومية وجمعيات النفع العام (ORG) فى مقدمة أنواع المواقع بنسبة تبلغ 46.6%تليها المواقع التعليمية (EDU) بنسبة 33.3% ثم المواقع التجارية بعد ذلك، وقد يشير ذلك إلى أهمية استخدام الجمل المصاغة بشكل فعال للابتعاد عن المواقع التجارية ومواقع الشركات وغيرها.

نتائج البحث المقارن بين المحركات الست :

قمـــنا بقـــياس تكرارية المواقع التي ظهرت كنتيجة للبحث، كما قمنا أيضاً بقـــياس ترتيـــب هذه المواقع في المحركات الستة، **وذلك بهدف الإجابة على سؤال محـــدد هـــو لمـــاذا تظهر هذه المواقع بالذات في هذه المحركات عند البحث عن مصطلح معين؟ وما هو مدى التشابه والاختلاف بين هذه المحركات؟**

وربمـــا ينبغـــي الإشارة إلى أن المحركات تعتمد في تكشيفها للمواقع على عنصرين:

1- العنصـــر الأول وهـــو المصطلحات التي وضعها الموقع في صفحة النص الفائق HTML والتي توضع في تركيبة داخلية في الصفحة تسمى Meta tag. هذه التركيبة تسمح بوضع مجموعة من المصطلحات التي تعبر عن المحـــتوى الموضوعي والنوعي للموقع، فعلى سبيل المثال عند الرغبة في وضع مصطلحات لموقع مكتبة عامة فإن هذه التركيبة تأخذ الشكل التالي:

page, collections and services, Thomas, copyright office,

هـــذه التركيـــبة تسمح لمحركات البحث باستخلاص الكلمات المفتاحية التي أشارت إليها الصفحة الأولى Home page لمكتبة الكونجرس.

2- العنصـــــر الثاني وهو المصطلحات التي قام محرك البحث باستخلاصها من الصـــفحة الأولى أو غيرها من الصفحات، وكما سبقت الإشارة فإن قدرات محـــركات البحث تختلف من محرك إلى آخر، فقد يقوم محرك باستخلاص الكلمـــات المفتاحـــية مـــن الصـــفحة الأولـــى فقط، وقد يقوم محرك آخر باستخلاص الكلمات المفتاحية من أكثر من صفحة.

إلا أن هـــناك ملاحظة عامة على العديد من المحركات أنها تقوم فى بعض الأحـــيان بوضع نتائج لمواقع - يلاحظ عليها أنها- قد لا تحتوى على المصطلح المـــبحوث عـــنه، وإنمـــا قـــد تكون هذه المواقع أقرب فى مفهومها وموضوعاتها للمصطلح، وهنا تبرز قضية هامة.. **ماهى الأدوات والسبل التى سلكتها المحركات لوضـــع هذه المواقع فى نتيجة البحث؟** وهذه القضية تحتاج لبحث أشمل وإن كانت هناك بعض الإجابات فى الأدب المنشور فى المجال.

جدول (8) المواقع المتشابهة والمختلفة والمكررة كنتيجة للبحث عن مصطلـــــــــح (Library) في المحركات الستة

م	المحرك	المواقع المتشابهة في المحركات الأخرى		المواقع المختلفة عن المحركات الأخرى		عدد المواقع المكررة في نفس البحث	
1	N.L.	صفر	.	10	%100	–	
2	Excite	4	%40	6	%60	–	
3	Google	8	%80	2	%20	–	
4	Ly cos	1	%10	9	%40	–	
5	All the web	3	%30	7	%7	–	
6	Alta vista	9	%90	1	%10	6	
	المجموع	25	%41.7	35	%58.3	%10	

تشير إحصائيات الجدول (8) إلى :

1- أن محـــرك Northernlight لم تتشابه نتائجه مع أي محرك آخـــر وبنسبة 100%، يلـــيه محرك لايكوس بنسبة 90%، ثم محرك all the web بنسبة 70%، ثـــم محـــرك Excite بنسبة 60%. ويمكن إرجاع ذلك إلى أن أغلب المحـــركات التـــي تشـــابهت نتائجها تعتمد في عرض نتائجها على الكلمات المفتاحية والواصفات التي وردت في تركيبة Meta tag في الصفحة الأولى لكـــل موقـــع، إضـــافة إلى إمكانية أن نقوم بإحصاء عدد مرات تكرار الكلمة

المـــبحوث عنها في بقية الصفحة أو الصفحات الأخرى التي قامت بعرضها. وفيما يتعلق بمحركي Northernlight و Lycos فإنهما يحتمل أنهما يعتمدان فـــي عرض نتائجهما على وجود نوع من التكشيف اليدوي[1] أو شبه اليدوي للمواقـــع، أو أنهما قد يعتمدان على عدد تكرارات الكلمة في الموقع، وبالتالي فـــإن ذلـــك لا يسمح بظهور نفس المواقع التي ظهرت كنتائج في المحركات الأخرى، إضافة إلى أنه عادة ما تقوم المحركات بوضع سياسة لعرض نتائج بحـــثها، قد تعتمد على عرض نتائج دولة معينة في البداية، أو مواقع بعينها- وفقاً للسياسة الإعلانية في المحرك - مما يحد من عرض مواقع قد تكون في غايـــة الأهمـــية بالنسبة للمستفيد. أو قد تلجأ بعض المحركات لاعتماد أسبقية تـــاريخ إنشاء الموقع فى الترتيب، وهنا يبرز السؤال، هل يتدخل تاريخ إنشاء الموقـــع فـــى تحديد ترتيبه فى قائمة نتائج البحث على أي محرك على شبكة

(1) تشير دراسة تمت فى جامعة البانى بالولايات المتحدة إلى أن محرك Northernlight يستخدم نوعا من الكشافات التى يتم معالجتها بشكل يدوى، إضافة إلى تكرار نفس النتائج فى المحركات الأخرى، وإلى أن محرك google يعتمد على استخراج نتائجه بناء على سلوك المستفيدين مع النتائج التى تعرض عليهم. راجع فى ذلك:
Second Generation Searching on the web. University of Albany. Retrieved and Printed. 15/9/2001. url < http://library.albany.edu/internet/second.html >

الإنترنـــت؟! وقد يحتاج هذا السؤال عند الإجابة عليه إلى مزيد من الدراسات والأبحاث لإثبات مدى صحته أو عدم صحته.

ورغــم ذلك فمن تحليل بنية الصفحة الأولى في المواقع التي ظهرت نتيجة للبحث، وجــد أنها جميعاً تحتــوي على تــركيبة Meta tag، ولكـــن عند بحـــث المحـــركات نفســـها وجـــد أن محـــرك Google لم يعتمد على هذه التركيبة في اســـتخلاص نتائجه فقط على الرغم من تشابه نتائجه بنسبة عالية بلغت 80% مع المحركات الأخرى.

2- يلاحـــظ أيضـــاً أن محرك Altavista تكررت نتائجه في نفس البحث بنسبة عالـــية. فقط ظهر موقع مكتبة الكونجرس 4 مرات في نفس البحث فيما ظهر موقـــع المكتـــبة الطبية الأميركية مرتين في نفس البحث، ومعنى ذلك أن ستة مواقـــع فقط هى التي احتوتها النتيجة، وليس عشرة مواقع، وربما يبرر ذلك النـــتائج الكبيرة التي ظهرت في عدد المواقع لهذا المحرك، أيضا يلاحظ هنا نـــوع مـــن الحشو فى قائمة النتائج وهو ما قد يفسر عدد النتائج الكبير الذي يخرج بها هذا المحرك.

3- كذلك ظهر موقع مكتبة الكونجرس كأول موقع في محركات ثلاثة هى محرك All the web ومحرك Excite ومحرك Google مما يشير إلى أهمية هذا الموقـــع بالنســـبة للمحـــركات الثلاثة، مما قد يعني من ناحية أخرى قيام هذه المحـــركات بعـــرض نـــتائج المواقـــع الأميركـــية أولاً، أو أنها تعتبر مكتبة الكونجرس الأميركية تستحق أن تعرض أولاً، أو أن هناك سياسة إعلامية، أو إعلانية تتعلق بهذه المكتبة في المحركات الثلاثة.

4- احتلت بعض المواقع الصدارة الأولى في المحركات التي تم البحث فيها وهي علـــى التوالي مكتبة الكونجرس حيث ظهرت سبع مرات في أربع محركات، والمكتـــبة الطبـــية الوطنـــية ظهرت أربع مرات في ثلاثة محركات، ومكتبة الإنترنت العامة والمكتبة البريطانية ظهرت ثلاث مرات في ثلاثة محركات.

ويبيـــن جدول (9) ترتيب المواقع التي ظهرت كنتائج لهذا البحث وفق عدد النقاط التي حصل عليها كل موقع في نتائج البحث (*).

(*) تم وضع نظام لحصول كل موقع على عدد معين من النقاط وفق الآتي :

1-ترتيبه بين نتائج البحث.

2-عدد مرات ظهوره كنتيجة للبحث في المحركات الست بحيث يحصل كل محرك على النقاط التالية :

عرض نتائج البحث	عدد النقاط
الموقع الأول	10 نقطة
الموقع الثاني	9 نقطة

جدول(9) المواقع التي ظهرت في الأبحاث الستة عن مصطلح (Library) وترتيبها

عدد النقاط	عدد مرات الظهور	الموقع
46	7	**1-Library of congress**
25	4	**2-National Library of medicine**
21	3	**3-Internet Public Library**
15	3	**4-The British Library**
14	2	**5-www.virtual Library**
9	2	**6-Lib web**
9	2	**7-ALA**
8	2	**8-New York Public Library**
7	2	**9-Electric Library**
4		**10-Christian classics ethereal Library**

وقد تم استبعاد المواقع التي وجدت مرة واحدة فقط أو المواقع التي لا تمت للموضوع المبحوث عنه بصلة (من وجهة نظر الباحث).

وبناء على جدول (9) فإن موقع مكتبة الكونجرس الأمريكية يحتل مقدمة المواقع التي حصلت على أفضل ترتيب، حيث حصلت على 46 نقطة يليها موقع المكتبة الطبية الوطنية الأمريكية، ثم موقع مكتبة الإنترنت العامة، وتساوى معه موقع المكتبة البريطانية، إلا أنه أتى تاليا له من حيث عدد النقاط بسبب موقفه فى قائمة ترتيب النتائج فى الأبحاث الستة التى تمت.

المواقع الثالث 8 نقطة
الموقع العاشر 1 نقطة

جدول (10) المواقع المتشابهة والمختلفة والمكررة كنتيجة للبحث عن مصطلح Digital Library في المحركات الستة

م	المحرك	المواقع المتشابهة في المحركات الأخرى		المواقع المختلفة عن المحركات الأخرى		عدد المواقع المكررة في نفس البحث	
		ع	%	ع	%	ع	%
1	N.L.	0	0	10	100	–	–
2	Excite	8	80	2	20	2	25
3	Google	6	60	4	40	2	33.3
4	Lycos	1	10	9	90	–	–
5	All the web	5	50	5	50	–	–
6	Alta vista	4	40	6	60	2	%50
	المجموع	24	40	36	60	6	10

بناء على جدول (10) يلاحظ ما يلي :

1- عـــدم تشابه نتائج المحرك N.L مع كل المحركات الأخرى، وبالاطلاع على نتائج هذا المحرك وجد أنها على الرغم من منطقياتها – من وجهة نظرنا– مع المصــطلح الــذي تم البحث عنه، فإن المحرك لم يأت بهذه المواقع من واقع استخدامه للكلمات والمصطلحات التي وردت في تركيبة Meta tag وإنما – الأحرى – أنه تم تكشيف هذه المواقع بشكل يدوي أو شبه يدوي، أو أنه يعتمد أسـلوبا فـــى ترتيب النتائج مختلفا عن المحركات الأخرى. ويؤكد ذلك تقسيم

نتائج المواقع والصفحات بناء على نوعية مصدر المعلومات، كأن يكون الناتج مقـــال، أو دليل، أو قائمة، أو معلومات عامة، إضافة إلى أنه يصنع نوعاً من التقيـــيم لكل موقع ناتج عن البحث بناء على عدد تكرارات المصطلح به، كما يؤكـــده مرة أخرى الاختلاف الكامل فى النتائج فى البحث الأول الذي تم على مصطلح Library.

2- كمـــا يأتـــي محـــرك Lycos بعد محرك .N.L في ضعف اتفاق نتائجه مع المحركات الأخرى حيث وجد به موقع واحد فقط متشابه مع نتائج المحركات الأخرى.

3- يتميز محرك Excite بتشابه نتائجه بنسبة عـــالية تصل إلى 80% من النتائج مـــع المحـــركات الأخـــرى، ويشـــير ذلك إلى ثبات هذا المحرك في عرض واستخلاص نتائجه من تركيبة Meta tag ومن محتويات صفحات الموقع.

4- يلاحـــظ أنه في نتائج البحث تكرر نفس العدد من المواقع في ثلاثة محركات هـــي Excite وGoogle وAltavista، ولكن بنسبة تختلف إلى العدد الكلي مـــن النـــتائج لكـــل محرك ففي محرك Excite مثلت 25% من النتائج وفي Google مثلت 33.3% من النتائج وفي Altavista مثلت 50% من النتائج نسبة إلى عدد المواقع المتشابهة مع المحركات الأخرى.

ويلاحظ هنا هبوط نسبة وعدد الصفحات المكررة في محرك Altavista بالتكرارات التي حققها في البحث السابق والذي تم على مصطلح واحد فقط (جدول (8)).

جدول (11) المواقع التي ظهرت في المحركات الستة عند البحث بمصطلح Digital Library وترتيبها

الموقع	عدد مرات الظهور	عدد النقاط
1-Berkeley	4	35
2-California Digital Library	7	49
3-N.Y.Public Library	3	20
4-ACM Digital Library	2	7
5-Networked Digital Library	4	22
6-Digital Library Federation	3	15

حيث احتل موقع Berkeley قمة الجدول بعدد 49 نقطة يليه موقع California Digital Library بعدد 35 نقطة، وإن كان الموقع الأخير قد أتى في الترتيب أولاً، في كل من محركي Excite وGoogle وهو ما يؤكد تماثل ترتيب النتائج في كل من المحركين حيث أتى أيضاً موقع Berkeley ثانياً في كليهما ثم اختلفت النتائج بعد ذلك بينهما.

جدول (12) المواقع المتشابهة والمختلفة والمكررة كنتيجة للبحـــث عن مصطلح (Digital Library) في المحركات الستة

م	المحرك	المواقع المتشابهة في المحركات الأخرى		المواقع المختلفة عن المحركات الأخرى		عدد المواقع المكررة في نفس البحث		ملاحظات
1	N.L.	1	10	9	90	–	–	
2	Excite	6	60	4	40	–	–	
3	Google	7	70	3	30	2	28.6	
4	Ly cos	9	90	10	10	–	–	
5	All the web	9	90	1	10	–	–	
6	Alta vista	3	30	7	70	2	66.6	
	المجموع	35	58.3	25	41.7	4	6.7	

يلاحظ على جدول (10) ما يلي :

1- ارتفاع معدل المواقع المتشابهة الناتجة عن البحث باستخدام مصطلحين معاً عنه فى البحث باستخدام مصطلح واحد، حيث انعكست هنا النتيجة. فقد كان عـــدد المواقـــع المتشابهة في البحث عن مصطلح واحد يشير إلى 41.7% بيـــنما ارتفـــع هـــنا إلـــى 58.3%، مما قد يشير إلى أنه كلما زادت عدد المصـــطلحات التـــي يـــتم البحث بها كانت هناك فرصة أكبر لتشابه نتائج البحـــث فـــي المحركات الستة، وذلك يشير أيضاً إلى ارتفاع درجة مطابقة المواقع للمصطلحات التي تم البحث بها.

2- يلاحـــظ أيضاً ارتفاع درجة مطابقة النتائج بين محركي Ly cos ومحرك All the Web وبنســـبة 90% من نتائج البحث، يليهما محرك Google وبنسبة 70% من النتائج، ثم محرك Excite وبنسبة 60% من النتائج. كما أن محـــرك N.1 مازال مختلفاً في نتائجـــه عـــن المحركات الباقية وبنسبة 90% وقـــد حقق نفس النتيجة في البحث بمصطلح واحد تقريباً، وهو ماقد يشير إلى تفرد هذا المحرك عن المحركات الأخرى – على الرغم من دقة نتائجه وتنوعها سواء من ناحية جنسية أو أنواع المواقع أو مضمونها – مما يعـــنى الحاجة إلي الدراسة المستفيضة لمحرك *NorthernLight* وسياسته فى عرض المواقع والأساليب والأدوات التى يسلكها فى هذا السبيل.

جدول (13) المواقع التي ظهرت في المحركات الستة عند البحث بمصطلح

Digital Library Collection

عدد النقاط	عدد مرات الظهور	الموقع
40	4	**-1New york Public Library**
30	4	**-2Digital Collection**
32	4	**-3American Memory**
14	2	**-4Networked Digital Library**
40	5	**-5Conell Digital Library**
14	3	**-6Making of America**
29	6	**-7Virginia Digital Library**
14	2	**-8NASD.L.**
8	2	**-9Americom Indiana**
9	3	**-10California D.L.**
6	3	**-11UCB Earth SCI**

حيث يلاحظ تساوي كل من موقعي N.Y.P.C وD.L وCornell وإن كان الموقـــع الأول قد ظهر في أربع محركات كأول موقع في الترتيب. والموقع الثاني ظهـــر خمـــس مـــرات، ولكن في أسبقيات مختلفة من محرك لآخر، وإن ظهر في Altavista وGoogle في الترتيب الخامس وظهر في كل من Lycos و All the web وNorthern Light فـــي الترتيب الثاني وتكرر مرتين في موقع All the web.

أما الموقع الأكثر ظهوراً فكان موقع Virginia Digital Library حيث ظهـــر ســـت مرات وإن تكرر مرتين في كل من Excite وLycos والغريب أن موقعي California وUCB أتيا في المركز الثامن والتاسع على التوالي في نتائج كل من محركات Excite و All the web وLycos. مما قد يعني مجموعة من الاســـتنتاجات حول تشابه التكنولوجيا المستخدمة فى المحركات والفروقات الزمنية فـــى تـــاريخ إنشـــاء هذه المحركات، والفروقات الناتجة من حيث تردد المصطلح، وتاريخ إنشاء المواقع نفسها، والمصطلحات الواردة فى تركيبة Meta Tag..

نتائج دراسة محركات البحث السالفة:

لوحظ من خلال التجربة المعملية التى أشرنا إليها أن هناك مجموعتين من النـــتائج، نتائج ذات طبيعة عامة خاصة بالمحركات كلها، ونتائج ذات طبيعة تتعلق بكل محرك على حدى، يمكننا الإشارة إليها فيما يلي:

1- النتائج العامة :

أ- يمكن القول بأن هناك مجموعة من العناصر التي تتحكم في نتائج البحث على أي محـــرك على شبكة الإنترنت، هذه العناصر بعضها يتعلق بالمحرك نفسه وقدراته وهي :

1- مدى دعم المحرك لصفحات الإطارات Frames وقدرته على تكشيفها.

2- مدى دعم المحرك لفقرات التعليقات Comments وقدرته على تكشيفها.

3- مـــدى قدرة المحرك على البحث في تركيبة Meta tag في صفحات المواقع وقدرته على استخلاص النتائج منها.

4- التكنولوجيا التي تم بناء المحرك بها وقدرتها على استخلاص النتائج المرجوة من صفحات المواقع أيا كان مكانها وموقعها.

5- مدى قيام المحرك بتنقيح الكشاف الذي يحتوي المواقع، وبالتالي إزالة المواقع (الميتة)، أو تحديث الروابط الموجودة بالفعل، أو إزالة الروابط المكسورة.

6- عدد الصفحات التي يقوم المحرك بتكشيفها في الموقع.

7- المدى الزمني الذي يقوم فيه المحرك (العنكبوت) بزيارة المواقع وتكشيفها.

8- الأهمـــية التـــي يوليها المحرك لتاريخ إنشاء الموقع وبالتالي ترتيبه بين نتائج البحث.

9- السياسة الإعلانية للشركة مالكة المحرك ومدى تأثيرها في ترتيب نتائج البحث.

10- استخدام نوع من التكشيف اليدوي فى المحرك، مما يعطى نتائج أدق من التكشيف الآلي.

11- استخدام أسلوب التعرف على نتائج البحث التى تمت من قبل، وقياس المواقع التى توجه إليها بالفعل المستفيدون من هذه النتائج.

وهناك **مجموعة من العناصر تتعلق بالمواقع نفسها التي يقوم المحرك بتكشيفها مثل :**

1- مدى ورود المصطلح من عدمه في تركيبة Meta tag.

2- عدد تكرارات المصطلح في تركيبة Meta tag.

3- عدد تكرارات المصطلح في الصفحة الأولى / أو الصفحات الأولى من الموقع.

4- عملية تصميم المواقع نفسها ومدى سماح المصمم للمحرك بتكشيف بعض الصفحات أو عدم تمكينه من تكشيفها.

وهناك مجموعة **من العناصر المتعلقة بالأجهزة المستخدمة (العميل والخادم) وقدرات الاتصال** وهي:

1- مواصفات الجهاز الخادم الذي سيتم الاتصال عن طريقه، فكلما كانت مواصفاته مرتفعة ومتقدمة كلما كان زمن البحث أسرع.

2- مواصفات الجهاز العميل الذي سيقوم بالاتصال، وينطبق عليه ما ينطبق على الجهاز الخادم.

3- خـط الاتصـال الذي يتم الاتصال عن طريقه، حيث تؤثر هذه القدرات على سرعة إنجاز البحث وهل يعمل الجهاز الخادم على خط مؤجر Leased Line ومعرف IP حقيقي أم من النوع البروكسي مما يساعد على ضعف أو قوة نقل البيانات وبالتالي سرعتها.

كذلك هناك مجموعة من العوامل المتعلقة بالمتصفح نفسه الذي سيستخدم في عرض البيانات. وكما هو معروف هناك متصفحان تجاريان يعملان على مستوى العالم في بيئة الحاسبات الشخصية هما متصفح إكسبلورر Explorer ومتصفح نت سكيب Netscape، والعناصر المتعلقة بالمتصفح هنا هي :

1- مدى دعم المتصفح لعرض الصفحات الناتجة عن البحث.

2- مدى دعم المتصفح لعرض الإطارات Frames.

3- مدى دعم المتصفح لعرض محرر نصوص HTML وإمكاناته في قراءة هذه النصوص.

4- سرعة المتصفح نفسه في عرض نتائج البحث (أو العامل الزمني في البحث).

تشكل هذه المجموعة من العناصر مجتمعة المؤثر الأساسي على سرعة عرض النتائج وطريقة الحصول عليها وبالتالي عرضها.

وبالنسبة للنتائج المقارنة يمكن الإشارة إلى مايلى:

1. لوحظ أن محرك Alltheweb هو صاحب أقل زمن بحثي بين المحركات الستة، وعلى الرغم من أن محرك Google قام برصد الزمن الذي تم فيه البحـــث ،إلا أن تدخل العوامل التي ذكرها الباحث تؤثر على زمن عرض النتائج على الجهاز العميل لديه، إضافة إلى عوامل أخرى غائبة مثل المدى الزمني الذي يتم فيه تنقيح كشاف المحرك، وتأثير تصميم المواقع نفسها فى ظهورها كناتج للبحث ومدى قدرة المحرك على تكشيف صفحات الإطارات والتعليقات، ومدى الالتزام بتكشيف تركيبة Meta Tag إلى جانب العوامل الأخرى.

2. لوحظ أيضا أن المحركات الخمسة الباقية تأتي في الترتيب كالتالي:

   1) محرك Lycos بمتوسط 4.3ث
   2) محرك Altavista بمتوسط 6.33ث
   3) محرك Excite بمتوسط 9.6ث
   4) محرك Northernlight بمتوسط 10.7ث

3. لوحـــظ أيضا أن محرك Altavista هو صاحب أعلى نسبة نتائج لإجمالي الاستفســـار يلـــيه محـــرك Google ثـــم محـــرك Excite ثـــم محرك Alltheweb ثم محرك Northernlight ثم محرك Lycos .

4. بالنسبة للبحث بكلمة مفردة كان صاحب أعلى عدد من الواقعات محرك Altavista وكان هو أيضا صاحب أعلى عدد من الواقعات في البحث بكلمتين مفردتين أو بثلاث كلمات.

5. لوحظ أن المحرك Northernlight صاحب أقل عدد من الواقعات، وأن تبادل المركز مع محرك Lycos.

6. لوحظ هيمنة المواقع الأمريكية التي تظهر كنتائج للبحث وبنسبة تراوحت على المحركات الستة بالنسبة للأبحاث الثلاث بين 85% :96.6%.

7. لوحظ أيضا أنه بالنسبة لمجموع الاستفسارات الثلاثة علي المحركات الستة جاء مواقع المنظمات(org.) في المرتبة الأولى بعدد 51 موقعا، ثم المواقع التعليمية (edu.) بعدد 49 موقعا، ثم مواقع الشركات بعدد 43 موقعا، ثم المواقع الحكومية بعدد 24 موقعا، ثم المواقع الأخرى مثل (net.)، أو تلك التابعة لدول لا يمكن تحديد هويتها بعدد 11 موقعا.

8. يلاحظ أيضا أنه عند البحث بكلمة مفردة واحدة أتت مواقع الشركات في المرتبة الأولى (جدول 3)، وربما يعود السبب في ذلك إلى اتساع استراتيجية البحث، بينما أتت مواقع المنظمات والمواقع التعليمية بأعلى نسبة نتائج عند البحث بجملة (جدول 7)، وربما يعود السبب في ذلك إلى ضيق استراتيجية البحث، وهو ما قد يميز هذا النوع من الأبحاث.

النتائج الخاصة بكل محرك:

وهي مجموعة النتائج المتعلقة بكل محرك بحث على حدى.

1- محرك البحث All the web (Fast):

أ – عدد الصفحات الناتجة: يلاحظ على هذا المحرك أنه كلما تم تضييق استراتيجية البحث كلما قل عدد المواقع وبالتالي الصفحات الناتجة وفقاً للنسب التالية :

متوسط عدد المواقع التى يعرضها كل محرك	البحث الأول	البحث الثانى	البحث بجملة
لكل 100 موقع ناتج	100	13	0.6

بمعنى أنه نسبة إلى كل 100 موقع في البحث بكلمة مفردة يتم عرض 13 موقعـــاً عند البحث بكلمتين، ويتم عرض 0.6 من الموقع عند البحث بجملة، وتتفق هـــذه النتيجة مع استراتيجيات البحث إجمالاً في أنه كلما ضاقت استراتيجية البحث كلها كان عدد النتائج أقل ومطابقاً إلى حد ما، عندما تم البحث عنه، ولكن لم يمكن دراسة الحشو Redundancy في الناتج المعروض – إلا فى أضيق الحدود– حيث يحـــتاج ذلـــك إلى بحث مواز يكشف عن مدى علاقة المواقع الناتجة بما تم البحث عنه، ويمكن القول بأن المواقع الناتجة احتوت على كلمات البحث التي تم البحث بها في الحالات الثلاث.

ب- حقــق هذا المحرك أقل زمن بحثي بين المحركات الستة، مما يؤكد على أهمية اســتخدامه في الأبحاث التي تتطلب سرعة الإنجاز، حيث حصل على متوسط 2.7 ث بالنســبة لحــالات البحــث الــثلاث. ويمكننا الزعم أيضاً بأنه يمكن اســتخدامه عــند العمــل على خطوط اتصال Dial up لبطء هذه الوسيلة في الاتصال، أو على خطوط مؤجرة ذات سرعات نقل ضعيفة لسرعته في عرض النتائج

جـ- فيما يتعلق بجنسية المواقع المعروضة أتت المواقع الأمريكية على رأس هذه المواقــع من ناحية الجنسية، وبمتوسط عدد وصل إلى 9.33 مواقع أمريكية و 0.33 موقعــاً بريطانــياً و0.33 موقعاً استرالياً. وتعود ظاهرة ارتفاع المواقع الأمريكــية إلى سيطرة هذه المواقع على شبكة الإنترنت، إضافة إلى أن اللغة الإنجليزية هى لغة أغلب مواقع الإنترنت. كما أن المصطلح الذي تم البحث به هــو مصطلح إنجليزي في الأساس، ولم يتم استخدام البحث بالمصطلح في كل اللغات، وهي إحدى خواص هذا المحرك.

د- فيما يتعلق بنوع المواقع أتت مواقع المنظمات في المرتبة الأولى بمتوسط 4.33 مواقــع ثم مواقع الشركات بمتوسط 2.66 موقعاً لكل عشرة مواقع، ثم المواقع التعليمــية بمتوسط 1.66 موقع لكل عشرة مواقع. وهذا المؤشر يعد مقبولاً -

مـــن وجهـــة نظـــر الباحث – لقياس مدى التنوع في عرض مواقع ذات فئة معينة.

وبصفة عامة يعد هذا المحرك مناسباً للأغراض البحثية والمهنية والتجارية إلـــى حد ما، ويمكن اللجوء إليه للأبحاث التي يجب إنجازها سريعاً خاصة في ظل اســـتجابته السريعة للأسئلة والاستفسارات المرجعية، كما يناسب الباحثين باستخدام خطوط Dial up.

2- محرك البحث Altavista

أ- عـــدد الصـــفحات الناتجة: يلاحظ على هذا المحرك أنه صاحب أعلى النتائج، والغريـــب أنه عندما تم تطبيق استراتيجية البحث (البحث الثالث) كانت النتيجة مـــرتفعة عن البحث بمصطلح واحد، إذ وصل عدد الصفحات التي استرجعها إلى حوالي 59 مليون صفحة وموقع، مما يؤكد أنـــه استخدام المعامل البولياني (OR) ولـــيس (And) وهذه النتيجة تؤكد على أهمية التعامل بحرص مع هذا المحرك وأهمية التأكيد على استخدام المعاملات البوليانية بشكل صحيح ووفقاً لرغبة المستفيد نفسه.

ب- حقق هذا المحرك زمناً بحثياً متوسطاً بلغ حوالي 6.33 ث وقد أتى في المرتبة الثالـــثة من حيث الزمن البحثي، على الرغم من النتائج الضخمة (عددياً) التي

قدمهـــا، ويمكن القول – بشكل نظري – أنه أسرع من محرك All the web قياساً إلى عدد نتائجه.

جـــ– فيما يتعلق بجنسية المواقع فقد أتت المواقع الأمريكية في المركز الأول 9.33 مواقـــع لكـــل عشرة مواقع و0.33 موقعاً بريطانياً لكل عشرة مواقع، و0.33 موقعاً استرالياً لكل عشرة مواقع، ويصدق عليه نفس القول الذي سبق الإشارة إليه في محرك All the web.

د– فيما يتعلق بنوعية المواقع أتت مواقع الشركات في المرتبة الأولى بعدد 4 مواقع لكـــل عشرة مواقع وأتت المواقع الحكومية في المركز الثاني 2.66 موقع لكل عشرة مواقع، وتساوت المواقع التعليمية والمنظمات في نفس العدد 1.66 موقع لكل عشرة مواقع، وعلى ذلك يمكن استخدام هذا المحرك عند البحث في مواقع الشـــركات، حيث بلغت نسبة هذه المواقع حوالي 40% من نسبة المواقع التي يقدمها.

وبشـــكل عـــام يمكـــن القول بمناسبة هذا المحرك للأغراض ذات الطابع الـــتجاري، وإن كانت ضخامة (عدد) نتائجه تؤكد أهميته في الأبحاث التي يندر أن تكون لها نتائج. وعلى ذلك يصدر – من وجهة نظر الباحث – باباً خلفياً للحصول على نتائج يمكن البحث فيها بعد ذلك خاصة للأكاديميين .

3- محرك البحث Excite

أ- عدد الصفحات الناتجة: بلغ متوسط عدد الصفحات الناتجة عن الأبحاث الثلاثة حوالـــي 6 ملـــيون صفحة وموقع، وهو يعد – من وجهة نظرنا– عدداً كبيراً نســبياً، كما يلاحظ تقارب نتيجة البحث الثاني والثالث مما قد يشير أيضاً إلى استخدام المحرك للمعامل البولياني (Or) بديلاً عن المعامل (And) مما يشير أيضاً إلى أهمية التعامل بحرص عند بناء استراتيجيات البحث والتعامل مع هذا المحرك، خاصة مع المعاملات البوليانية وأساليب البحث الأخرى.

ب- زمن البحث: حقق هذا المحرك زمناً كبيراً قياساً إلى المحركات الأخرى – مع أهمية القول بأن هذه المساحة الزمنية نسبية.

2- جنسية المواقع: أتت المواقع الأميركية في قمة الهرم بنسبة عالية حيث احتلت 9.66 لكل عشرة مواقع، وأتت معها المواقع البريطانية بعدد 0.33 موقعاً لكل عشرة مواقع مما يعني أهمية هذا المحرك للباحثين عن المواقع الأمريكية بشكل عام.

3- نوعـــية المواقع: أتت مواقع الشركات على قمة المواقع بعدد 4.33 مواقع لكل عشرة مواقع، مما يجعله يسبق محرك الطافيستا، ويؤكد على استخدام المحرك للباحثيـــن عـــن مواقـــع الشركات، خاصة في الاستفسارات التي تحمل الطابع التجاري.

وبصـــفة عامـــة يمكـــن القول بمناسبة هذا المحرك للأغراض ذات الطابع التجاري خاصة أن نوعية المواقع التي قدمها في أغلبها شركات .

4- محرك البحث Google

أ- عـــدد الصـــفحات الناتجة: بلغ متوسط الصفحات الناتجة عن الأبحاث الثلاثة حوالي 12.7 مليون صفحة وموقع، ولكن لاحظ أنه يستخدم المعامل البولياني (and) ولـــيـــس (or) بدليل أن البحث الثاني والثالث حقق نتائج قليلة للغاية قياســـاً إلى البحث الأول، مما يؤكد على أن استخدامه يتم بشكل صحيح خاصة بالنســبة للأكاديميين وراغبي البحث عن موضوعات محددة وحقيقة وسلامة نتائج البحث التي ستعرض.

ب- الزمن: حقق زمناً متوسطاً بلغ 7.8 ث بالنسبة للأبحاث الثلاثة وإن كان يلاحظ أن هـــناك تـــدرج زمني في نتائج كل من الأبحاث الثلاثة بلغ حوالي 1.5 ث بالزيادة بين كل بحث وآخر.

جـــــ- جنسية المواقع: أتت المواقع الأمريكية بعدد 8.66 موقعاً لكل 10 مواقع ثم المواقـــع البريطانـــية والنيوزيلندية بنفس النسبة، ولم يتم التعرف على جنسية موقعيـــن مـــن المواقـــع الـــناتجة عن البحث الأول، ويلاحظ هنا نوعاً – ولو ضئيلاً– في التغطية المكانية للمواقع.

د- نوعية المواقع: أتت المواقع التعليمية ومواقع المنظمات في المركز الأول بنفس العدد وهو 3.66 صفحة، وموقع لكل عشرة مواقع، مما يؤكد أيضاً على مناسبة هذا المحرك للأغراض الأكاديمية والبحثية وللباحثين عن النصوص الكاملة.

وبصفة عامة فإن طريقة البحث وناتج البحث سواء من نوعية المواقع وإلى حد ما من جنسيتها يؤكد على أهمية المحرك للأغراض البحثية والأكاديمية .

5- محرك البحث Lycos

أ- عدد الصفحات الناتجة: بلغ متوسط عدد الصفحات والمواقع الناتجة عن الأبحاث الثلاثة حوالي 6 مليون صفحة وموقع، كما أنه يستخدم مباشرة المعامل البولياني (and) مما يؤكد أيضاً على عمق النتائج المستخلصة، حيث يمكن الإشارة أيضاً إلى أن نتيجة البحث الثالث وصلت إلى 256 ألف صفحة وموقع، مما يشير أيضاً إلى أهميته بالنسبة للأكاديميين والباحثين العلميين حيث يقدم نتائج دقيقة إلى حد معين بعد مراجعة الباحث للعديد من المواقع الناتجة عن الأبحاث التي قام بها بشكل تجريبي.

ب- الزمن: حقق هذا المحرك ثاني أفضل نتيجة من حيث سرعة زمن البحث، إذ وصل إلى متوسط 4.3 ثانية مما يؤكد أيضاً على أهمية استخدامه في الأبحاث

التي تتطلب نوعاً من السرعة، أو للباحثين على خطوط من نوع Dial up، أو حتى الخطوط المؤجرة Lased Line المزدحمة أو الضعيفة.

جـ- جنسية الموقع: سيطرت المواقع الأميركية بشكل ساحق على نتائج هذا المحرك وبعدد 9.66 لكل عشرة مواقع، وأتت المواقع الأسترالية بعدد 0.33 لكل عشرة مواقع، مما يشير أيضاً بأن هذا المحرك يعكس الوجود والثقافة الأمريكية وهو يعتبر مناسباً للباحثين في هذا الاتجاه.

د- نوع المواقع: أتت المواقع التعليمية في المركز الأول بعدد 4 مواقع لكل عشرة مواقع، ثم مواقع الشركات، ثم مواقع المنظمات.

وعلى ذلك فهو مناسب للبحث في المؤسسات ذات الطابع التعليمي، خاصة من قبل الباحثين الأكاديميين، وفي ذات الوقت يناسب الباحثين في المؤسسات ذات الطابع التجاري.

وبشكل عام فهو صالح للباحثين الأكاديميين لدقة نتائجه وسرعته.

6- محرك البحث *Northernlight* :

أ- عدد الصفحات الناتجة: بلغ متوسط عدد الصفحات الناتجة عن الأبحاث الثلاثة 5 مليون صفحة، ويلاحظ أن البحث الثالث كانت نتيجته 305 ألف صفحة، مما يشير أيضاً إلى استخدام المحرك للمعامل (and). أيضا

ب- الزمـــن: حقـــق هذا المحرك متوسطاً زمنياً بلغ 10.7 ث للبحث الواحد وهو معدل زمني عال بالنسبة للمواقع الأخرى.

جـــــ- جنســـية المواقـــع: يعد الموقع الوحيد الذي شهد تنوعاً كبيراً في جنسيات المواقـــع، وعلى الرغم من السيطرة الطبيعية للمواقع الأمريكية إلا أنه ظهرت فـــيه مواقع من كل من استراليا وألمانيا وهولندا ونيوزيلندا وفرنسا، مما يؤكد على إيجابيته في استعراض جنسيات المواقع وما في ذلك من جاذبية للباحثين سواء الأفراد أو المؤسسات الراغبين في التنوع.

د- نـــوع المواقـــع: أتت المواقع التعليمية بأغلبية كبيرة وصلت إلى 4 مواقع لكل عشـــرة مواقـــع مما يؤكد مرة أخرى على مناسبته العالية للباحثين في المجال الأكاديمي.

وبصـــفة عامـــة ينفرد هذا المحرك بميزة قد لا توجد في أغلب المحركات الأخـــرى وهي أنه يقسم المواقع بالنسبة إلى الفئة المرجعية التي تنتمي إليها سواء كانت مراجع أو مصادر معلومات أخرى مثل المقالات والكتب وغيرها، مما يجعله أنســـب محـــرك بين المحركات الستة للباحثين في مصادر معلومات محددة، أو في مواقع ذات نوعية محددة، حيث أنه يقدر هذه الأفضلية.

النتائج المقارنة بين المحركات:

يمكننا الآن أن نضع بعض النتائج المقارنة بين المحركات لنتمكن فى النهاية كقراء من استخدام هذه المحركات بشكل جيد ومخطط. فإذا قورنت هذه النتائج بين المحركات الستة فإنه يلاحظ الآتي:

1- محركات NL و Lycos و Excite و All the web تسترجع جميعا بمتوسط 5 -6 مليون صفحة أو موقع للبحث الواحد فيما يشير كل من AltaVista و Google، مما يشير إلى أن المحرك المقبول الملائم يسترجع هذا العدد في البحث الواحد.

2- المدى الزمني الملائم للاسترجاع بين 2-6 ث لأي محرك، وهذا الزمن يعد مقبولاً من وجهة نظرنا، وإن كان يعتمد على مدى سعة كشاف المحرك، وعدد المواقع التي تمكن من تكشيفها وفهرستها.

3- أهمية التنوع في نتائج البحث من حيث جنسية المواقع التي يعرضها المحرك في نتائجه، فعلى الرغم من السيطرة الأمريكية على المواقع على شبكة الإنترنت، إلا أنه تظل هناك فرصة لظهور مواقع من جنسيات أخرى كما هو الحال في محرك N.L. .

4- تبرز أهمية شبكة الإنترنت بالنسبة للباحثين العلميين في كمية المواقع التعليمية ومواقع المنظمات والمواقع الحكومية التي تقدم نصوصاً وبيانات كاملة،

وتحـــاول محـــركات مـــثل N.L وLycos وGoogle أن توفر هذه المواقع بكميات كبيرة أمام هؤلاء الفئة من متصفحي الشبكة العنكبوتية العالمية.

يمكننا أيضا أن نشير فى عجالة إلى مجموعة من العناصر التى تؤثر على عملية فحـــص المواقـــع الـــتى يقوم بها كل محرك على الإنترنت، وهذه المجموعة من العوامل مترجمة ولكن بتصرف.

العوامل المؤثرة على عملية فحص المواقع :

1- **الحجـــم Size**: كلمـــا كان محرك البحث كبيراً – بناء على عدد الصفحات المكشـــفة – كلما كان عدد الصفحات التي يمكن تكشيفها في المواقع أكثر. ومـــن هنا يتم وضع محركات البحث في أربع فئات هى كبيرة، متوسطة، وصغيرة، وضئيلة.

2- **عـــدد الصـــفحات التي يتم فحصها يومياً**: يبين هذا العنصر إلى أي مدى يمكن لمحرك البحث أن يكشف عددا من الصفحات في اليوم الواحد. وكلما كانـــت قـــدرة المحرك على الفحص والتجول في المواقع كبيرة، كلما كان الكشـــاف الخـــاص بـــه أكـــثر حداثة وتعبيراً عن صحة البيانات، وبعض محركات البحث لديها القدرة على التعلم كيف يمكن للصفحات أن تتغير أو أنهـــا تســـتخدم طـــرقاً أخرى لتحسين عمليات التحديث لزيادة القدرة على فحص المواقع.

3- **الحداثـــة**: تتغير الشبكة العنكبوتية بشكل مستمر، وعلى ذلك فمن السهل أن تـتقادم المواقـــع المكشـــفة فـــي محركات البحث، وعلى ذلك فبعض هذه المحـــركات قـــد تكـــون قوائمها لم تحدث منذ يوم أو بعض يوم أو أنها لم تحدث منذ شهر أو يزيد.

وهـــناك العديد من الأسباب التي تقف وراء ذلك. فبعض محركات البحث تكشـــف (بالصـــدفة) أية صفحة تصل إليها، وتستغرق وقتاً أطول لفحصها والعودة وجمـــع الصـــفحات التـــي لـــم ترضخ لها، كما قد تقوم محركات البحث بتكشيف الصفحات الأكثر جماهيرية على الشبكة وتهمل الأجزاء الأخرى.

إن عنصر الحداثة بيِّن فى عمر القوائم التي تظهر، وبالتالي المدى الزمني أو مدى التقادم في تكشيف هذه المواقع والصفحات.

4- **الصـــفحات الراضخة**: إذا اتبعنا توالي العمليات بشكل منطقي، فإن محرك البحث سوف يعثر على الصفحات على الشبكة طالما يقوم بتتبع الوصلات والروابط Links أثناء قيامه بفحص Crawling الشبكة.

وفـــي الواقـــع، فإن الصفحة سوف تظهر بسرعة أكثر إذا خضعت مباشرة للمحرك. فيبين ذلك العامل كيف يمكن أن تتوقع أن خضوع صفحة للمحرك سوف يعمل على ظهورها في قوائم محرك البحث في الحال.

5- **الصفحات غير الخاضعة**: عند خضوع الصفحة للمحرك فإن محرك البحث سوف يجد صفحات أخرى عادة في الموقع تتبع الصفحة الخاضعة. وعلى أيـــة حال، فإن بعض المحركات تأخذ وقتاً أطول في جمع الصفحات غير الخاضعة. وعملياً، يحدث ذلك لسبب قيام بعض المحركات بتكشيف صفحة (بالصــدفة) تكــون خاضعة، ويضيف الموقع إلى جدول عمليات الفحص والـــزحف فـــي المســتقبل. ويبين الجدول السابق كيف يمكن أن نتوقع أن الصفحات الأخرى من الموقع سوف تظهر طالما تم خضوع صفحة واحدة مـــن الموقع، وبافتراض أنه ليس هناك مشكلات تمنع المحرك من العثور على هذه الصفحات مثل الإطارات أو الخرائط المصورة كما سيلي شرحها.

6- **العمق**: يرتبط هذا العامل بالصفحات غير الخاضعة للمحرك فهي تشير إلى عـــدد الصـــفحات التي تقف خلف الصفحة الخاضعة والتي سيقوم المحرك بجمعها، لعمل المحرك بأسلوبين أو طريقتين :

- أســـلوب المسح الشامل: بمعنى أن المحركات سوف تحاول جمع أي شيء تجـــده في الموقع، وقد لا تحصل على كل صفحة، ولكن سيظل ذلك دائما هو الهدف العام.

- أسلوب العينة: هذه النوعية من المحركات تجمع عينة من صفحات الموقع بعضها يجمع عينة أكبر من الأخرى، وباستخدام الحجم المحدد سلفاً كدليل

على حجم العينة، يمكن أن نتوقع عدد الصفحات التي قام المحرك بجمعها، عـــادة، كلمـــا كان الموقع عاماً وشائعاً، فمن المستحب عرضه من خلال المحـــرك. ويجب أن يكون معلوماً أن جزءاً من الشبكة يظل غير مكشف لاعتـــبارات عديدة، مثل الإطارات وخرائط الصور والصفحات التي يتم تخليقها آليا كلها عرضة لأن لا يعرضها المحرك.

7- **دعـــم الإطارات* Frames support**: هل يمكن لمحرك البحث أن يتتبع وصلات الإطارات؟ إذا كان باستطاعته فربما يفقد العديد من الصفحات في الموقـــع. فالصـــفحات الإطارية تتكون من صفحات منفصلة قد يكون من الصعب على المحرك أن يتتبعها جميعاً.

8- **الخـــرائط المصـــورة Image Maps**: هل يمكن لمحرك البحث أن يتبع الجانـــب الآخـــر للخرائط المصورة لحاسب العميل side–Client فمع الإطـــارات إذا لم يستطع محرك البحث تتبع الخرائط فمن المحتمل أن نفقد العديد من المعلومات عن الموقع.

(الخـــرائط المصورة هى صفحة تبين التنظيم الداخلي للموقع وما يرتبط به من وصلات. وعادة ما تكون هذه الصفحة على شكل صورة تم مسحها ضوئياً وتم

*الإطارات واحدة من خصائص لغات النص الفائق بحيث تمكن من تداخل أكثر من صفحة فى صفحة واحدة / وحيث أنها تظهر أمام المستفيد كأنها صفحة مقسمة إلى عدة أجزاء بينما هى فى الحقيقة عدة صفحات منفصلة. (المترجم)

إعدادهــا ببرنامج خاص بالرسم. وعلى ذلك فمن الصعب لمحرك البحث الحصول منها على معلومات حيث أنه يتعامل مع صورة في تلك الحالة).

9- **المواقــع المحمــية بكلمــات عبور**: بعض محركات البحث يمكنها دخول المواقع المحمية بكلمات عبور، إذا سهلت لها أن تحصل على اسم مستخدم username وكلمة عبور password – ولماذا تفعل ذلك؟ لربما أردنا من الناس اكتشاف أن لدينا في موقعنا مواد تطابق استفساراتهم، ورغم ذلك فهم مــازالوا بحاجة لملء بطاقة تسجيل معلومات مناسبة في موقعنا للوصول إليه، ولكن على الأقل سيعرفون أنه موجود.

10- **شيوع الرابطة أو الواصلة**: كل محركات البحث يمكنها تحديد مدى شيوع صفحة بتحليل عدد الروابط الموجودة بها. بعض المحركات تستخدم ذلك كوسيلة لتحديد أية صفحة سوف يتم تضمينها في الكشاف.

11- **الــتعلم بالــتردد على الصفحة**: عدد من محركات البحث يمكنها أن تتعلم وتعــرف إلـــى أي مـــدى يمكن لصفحة ما أن تتغير ومواقيت التغير هنا الموقع الذي يتغير غالباً ما يتم زيارته دائماً. هذه التغيرات المفاجئة تحتاج زيارات منظمة.

12- **معيار / ملف** Robots.text [1] هــذا الملف أو المعيــار يعنــي بالنسبة للمســئولين عن المواقع على الإنترنت webmasters أن تبقى محركات البحث بعيدة عن مواقعهم.

13- حقــل meta Robots: هــذا الحقل الخاص يسمح لمالكي الموقع بتحديد الصفحة التي لا يجب تكشيفها في الموقع، وهذا الحقل مثالي لهؤلاء الذي لا يســتطيعون بناء ملف robots.txt، ولإبقاء محركات البحث بعيداً عن الصفحة وبالتالي عدم تكشيفها يتم كتابة السطر التالي :

14- **اختبار حالة عنوان الموقع URL**: يبين ذلك ما إذا كانت صفحة الموقع قــد تــم تكشــيفها من قبل محرك البحث. "yes" تعني أنه يمكنك البحث بسهولة عن صفحة محددة وترى كيف تظهر بالضبط في الكشاف، ويعني هــذا الشــكل أن هــناك حالة اختبار لعنوان الموقع ستخبرنا ما إذا كانت الصفحة تم تكشيفها من عدمه، وكذلك لن ترى القائمة الفعلية.

هذه العوامل الأربعة عشر هى التي تؤثر في مدى ظهور الموقع في محرك بحث من عدمه.

---

[1] لمزيد من المعلومات يمكن زيارة الموقع التالي :
URL: http: // info.webcrawler.com/wak/projects/robots/exclasion.html

المراجع والمصادر

1. http://info.webcrawler.com/wak/projects/robots/exclasion.html
2. http://www.alltheweb.com/
3. http://www.altavista.com/
4. http://www. Beaucoup.com / engines.htm.
5. http://www.excite.com/
6. http://www.google.com/
7. http://www.happytrails.com / target.htm
8. http://www.lycos.com/
9. http://www.northernlight.com/
10. http://www.nueva.pvt.k12.ca.us/~debbie/library/research/adviceengine.html/
11. Page, Adam. The search Engines secrets of the pros. 9p. In: www.znet.com/Pccomp/features/fea1096/sub2.html (oct. 1998)
12. Catlege, Lara D. Characterizing browsing strategies in the World Wide Web. 12p. retrieved and printed: 2/7/2001 URL<http://www.igd.fhg.de/archive/1995_www95/papers

13.Peterson, Richard Einer. Eight Internet search engines compared. First Monday journal. 9p. Retrieved and printed 12/7/2001.
URL<http://www.firstmonday.dk/issues/issue2_2/peterson /index.html>

14.Sprink, Amanda, Bateman, Judy and Jansen, Bernard J. Searching the web: a survey of excites users. Internet research: electronic applications and policy. Vol. 9. No. 2. 1999. Pp: 117-128

15.Bar-ilan, Judit. Evaluating the stability of the search tools Hotbot and Snap: a case study. Online information review. Vol. 24. No. 6.

16.Platt, Nina. Website search engines. November 1996. the piper letter: Databases, Features and special reports.www.piperinfo.com/p103/search.html Last updated: Thu. Feb. 20,1997 retrieved and printed: 22-12-1998

17.how search engines work. In web site titled: search engine watch .url <http://www.searchenginewatch.com /webmasters/work.html>. Retrieved & printed: 18/10/1998.

18.How search engines rank web pages. IN:
http://www.go.com/

19.Nueva Library help. Choose the best engine for your purpose. (Oct. 1998)
www.nueva.pvt.k12.ca.us/~debbie/library/research/adviceengine.html/

20.Search Engine Tutorial for Web Designers: Northern Light. http://northernwebs.com/set/northern_lights.html

21.Search Engine Tutorial for Web Designers: AltaVista. url <http://northernwebs.com/set/alta_vista.html>

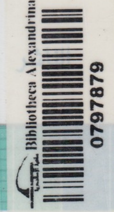
22.Search Engine Tutorial for Web Designers: Lycos. url <http://northernwebs.com/set/Lycos.html>

23.Second Generation Searching on the web. University of Albany. Retrieved and Printed. 15/9/2001. url < http://library.albany.edu/internet/second.html >

يعد هذا الكتاب واحد من أهم الكتب التي تتناول أدوات البحث على شبكة الإنترنت بشكل موسع وبسيط ليكون في متناول كل قارئ يهتم بالإنترنت ومجال تكنولوجيا المعلومات بشكل عام.

يتناول المؤلف مجموعة من أشهر أدوات البحث على شبكة الإنترنت وهى التي تمثل مفاتيح وخرائط للبحث في تلك الشبكة العالمية، إذ لايمكنك البحث عبر الشبكة دون أن تملك القدرة على التعامل مع هذه المفاتيح التي تتيح لك البحث عبر الإنترنت بكل سهولة ويسر، إضافة إلى أن المؤلف أضاف مجموعة رائعة من المعلومات حول طرق عمل محركات البحث وهو مايجب أن يفهمه كل باحث على الإنترنت.

يعتز الناشر بنشر هذا الكتاب المتفرد في موضوعه لواحد من أشهر الكتاب في العالم العربي في مجال تكنولوجيا المعلومات.



الطبعة الأولى 2007